

PATENT COOPERATION TREATY

PCT

NOTIFICATION OF ELECTION

(PCT Rule 61.2)

From the INTERNATIONAL BUREAU

To:

Assistant Commissioner for Patents
 United States Patent and Trademark
 Office
 Box PCT
 Washington, D.C. 20231
 ÉTATS-UNIS D'AMÉRIQUE

in its capacity as elected Office

Date of mailing (day/month/year) 13 August 1999 (13.08.99)	
International application No. PCT/IB98/02085	Applicant's or agent's file reference 339507/17829
International filing date (day/month/year) 04 December 1998 (04.12.98)	Priority date (day/month/year) 04 December 1997 (04.12.97)
Applicant LADANT, Daniel et al	

1. The designated Office is hereby notified of its election made:

☒ in the demand filed with the International Preliminary Examining Authority on:
 28 June 1999 (28.06.99)

☐ in a notice effecting later election filed with the International Bureau on:

2. The election ☒ was
☐ was not



made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).

The International Bureau of WIPO 34, chemin des Colombettes 1211 Geneva 20, Switzerland Facsimile No.: (41-22) 740.14.35	Authorized officer F. Zotomayor Telephone No.: (41-22) 338.83.38
---	--

PCT

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference 339507/17829		FOR FURTHER ACTION See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)	
International application No. PCT/IB98/02085	International filing date (day/month/year) 04/12/1998	Priority date (day/month/year) 04/12/1997	
International Patent Classification (IPC) or national classification and IPC G01N33/50			
Applicant INSTITUT PASTEUR et al.			
<p>1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.</p> <p>2. This REPORT consists of a total of 8 sheets, including this cover sheet.</p> <p><input type="checkbox"/> This report is also accompanied by ANNEXES, i.e. sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).</p> <p>These annexes consist of a total of sheets.</p>			
<p>3. This report contains indications relating to the following items:</p> <ul style="list-style-type: none">I <input checked="" type="checkbox"/> Basis of the reportII <input type="checkbox"/> PriorityIII <input checked="" type="checkbox"/> Non-establishment of opinion with regard to novelty, inventive step and industrial applicabilityIV <input type="checkbox"/> Lack of unity of inventionV <input checked="" type="checkbox"/> Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statementVI <input checked="" type="checkbox"/> Certain documents citedVII <input checked="" type="checkbox"/> Certain defects in the international applicationVIII <input checked="" type="checkbox"/> Certain observations on the international application			
Date of submission of the demand 28/06/1999		Date of completion of this report 23.02.00	
Name and mailing address of the international preliminary examining authority:  European Patent Office D-80298 Munich Tel. +49 89 2399 - 0 Tx: 523656 epmu d Fax: +49 89 2399 - 4465		Authorized officer Moreno de Vega, C Telephone No. +49 89 2399 7486 	

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No. PCT/IB98/02085

I. Basis of the report

1. This report has been drawn on the basis of (*substitute sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to the report since they do not contain amendments.*):

Description, pages:

1-43 as originally filed

Claims, No.:

1-45 as originally filed

Drawings, sheets:

1/7-7/7 as originally filed

2. The amendments have resulted in the cancellation of:

- ☐ the description, pages:
☐ the claims, Nos.:
☐ the drawings, sheets:

3. ☐ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed (Rule 70.2(c)):

4. Additional observations, if necessary:

III. Non-establishment of opinion with regard to novelty, inventive step and industrial applicability

The questions whether the claimed invention appears to be novel, to involve an inventive step (to be non-obvious), or to be industrially applicable have not been examined in respect of:

- ☐ the entire international application.
☒ claims Nos. 40-42.

because:

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No. PCT/IB98/02085

☐ the said international application, or the said claims Nos. relate to the following subject matter which does not require an international preliminary examination (*specify*):

☒ the description, claims or drawings (*indicate particular elements below*) or said claims Nos. 40-42 are so unclear that no meaningful opinion could be formed (*specify*):

see separate sheet

☐ the claims, or said claims Nos. are so inadequately supported by the description that no meaningful opinion could be formed.

☐ no international search report has been established for the said claims Nos. .

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Statement

Novelty (N)	Yes:	Claims 2-39, 43-45
	No:	Claims 1
Inventive step (IS)	Yes:	Claims 2-39, 43-45
	No:	Claims 1
Industrial applicability (IA)	Yes:	Claims 1-39, 43-45
	No:	Claims

2. Citations and explanations

see separate sheet

VI. Certain documents cited

1. Certain published documents (Rule 70.10)

and / or

2. Non-written disclosures (Rule 70.9)

see separate sheet

**INTERNATIONAL PRELIMINARY
EXAMINATION REPORT**

International application No. PCT/IB98/02085

VII. Certain defects in the international application

The following defects in the form or contents of the international application have been noted:

see separate sheet

VIII. Certain observations on the international application

The following observations on the clarity of the claims, description, and drawings or on the question whether the claims are fully supported by the description, are made:

see separate sheet

**INTERNATIONAL PRELIMINARY
EXAMINATION REPORT - SEPARATE SHEET**

International application No. PCT/IB98/02085

Reference is made to the following documents:

- D1: PELLETIER J N ET AL: 'A PROTEIN COMPLEMENTATION ASSAY FOR DETECTION OF PROTEIN-PROTEIN INTERACTIONS IN VIVO' PROTEIN ENGINEERING, vol. 10, no. SUPPL. 01, 1 October 1997, page 89 XP002064563
- D2: ROSSI F ET AL: 'MONITORING PROTEIN-PROTEIN INTERACTIONS IN INTACT EUKARYOTIC CELLS BY BETA-GALACTOSIDASE COMPLEMENTATION' PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 94, August 1997, pages 8405-8410, XP002064565 cited in the application

Re Item III

Non-establishment of opinion with regard to novelty, inventive step and industrial applicability

Claims 40-42 do not meet the requirements of Article 6 PCT in that the matter for which protection is sought is not clearly defined, i. e. any eucaryotic and prokaryotic protein, antibodies, cofactors, etc. falls within the scope of the claims. Besides, the wording of claim 41 is so unclear that it is difficult to interpret what is claimed.

Re Item V

Reasoned statement under Rule 66.2(a)(ii) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. D1 discloses a protein complementation assay for detection of protein-protein interactions in vivo, based on reconstitution of dihydrofolate reductase (DHFR) activity by complementation of defined fragments of the murine enzyme (mDHFR) in *E. coli* (see first paragraph and figures).
- D2 presents an approach for monitoring protein-protein interactions within intact eukaryotic cells, in which enzymatic β -gal activity serves to monitor the formation of the rapamycin-induced chimeric FRAP/FKBP12 protein complex

in a time-and-dose dependent manner, as assessed by histochemical, biochemical and fluorescence-activated cell sorting assays (see abstract, figure 1, page 8405, right column, paragraph 2 - page 8406, 1st. paragraph).

2. Claim 1 is not considered to be new (Article 33 (2) PCT), because D1 and D2 anticipate the signal amplification system claimed.

Claims 2-9 dependent on claim 1 contain features which in combination with claim 1 define a subject-matter which is considered to be new, i.e. the use in the signal amplification system of adenylate cyclase and guanylate cyclase as enzymes which trigger a signal amplification.

Claims 10-39 and 43-45 are considered to be new, the method of selecting a molecule on interest, the kit therefore, the method of screening for a substance capable of stimulating or inhibiting the interaction between a target ligand and a molecule of interest, the kit therefore, the signal amplification system and the polynucleotide sequences therefore claimed are not described in the prior art.

3. The technical problem to be solved by claims 2-9 and 43-45 is considered to be the provision of a signal amplification system to detect interactions of a molecule of interest and a target ligand; in claims 10-39 the technical problem to be solved is the provision of a method of selecting a molecule of interest capable of binding to a target ligand and a method of screening substances capable of stimulating or inhibiting the interaction between a target ligand and a molecule of interest.

Prior art documents D1 and D2 try to solve the same technical problem using as signal the enzymatic activity of DHFR and beta-galactosidase. The solution proposed by the present invention, using adenylate cyclase and guanylate cyclase as signalling molecules which trigger transcriptional activation is not disclosed nor suggested by the prior art, thus the subject matter of claims 2-39 and 43-45 is considered to involve an inventive step.

**INTERNATIONAL PRELIMINARY
EXAMINATION REPORT - SEPARATE SHEET**

International application No. PCT/IB98/02085

4. Should the priority of the present application not be valid, the following document would be relevant with respect to a novelty and inventive step (Articles 33(2) and (3) PCT).

KARIMOVA, GOUZEL ET AL: 'A bacterial two-hybrid system based on a reconstituted signal transduction pathway' PROC. NATL. ACAD. SCI. U. S. A. (1998), 95(10), 5752-5756 CODEN: PNASA6; ISSN: 0027-8424, 1998, XP002100623

Re Item VI

Certain documents cited

Certain published documents (Rule 70.10)

Application No Patent No	Publication date (day/month/year)	Filing date (day/month/year)	Priority date (valid claim) (day/month/year)
WO 98 34120	6.8.1998	2.2.1998	31.1.1997

Should the priority of the present application not be valid, the above document would be relevant with respect to novelty and inventive step (Articles 33(2) and (3) PCT). Furthermore, should the present application be entered into the regional phase, the above document could be relevant to the question of novelty.

Re Item VII

Certain defects in the international application

The incorporated prior art by reference is not allowed, as the application should be self-contained. Therefore, the sentence "...are incorporated by reference" on page 2 line 17 and page 19 line 30 should be deleted.

Contrary to the requirements of Rule 5.1(a)(ii) PCT, the relevant background art disclosed in the document D1 is not mentioned in the description, nor is this document

**INTERNATIONAL PRELIMINARY
EXAMINATION REPORT - SEPARATE SHEET**

International application No. PCT/IB98/02085

identified therein.

Re Item VIII

Certain observations on the international application

1. Claims 40 does not meet the requirements of Article 6 PCT in that the matter for which protection is sought is not clearly defined, the claim defining the subject-matter in terms of the result to be achieved.
2. The wording "in an E. coli strain, or in any bacterial strain deficient in endogenous adenylate cyclase or any other eukaryotic cell" in claims 22, 23, 24, 38, 39 is unclear, contrary to Article 6 PCT.

PCT

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference 339507/17829	FOR FURTHER ACTION see Notification of Transmittal of International Search Report (Form PCT/ISA/220) as well as, where applicable, item 5 below.	
International application No. PCT/IB 98/02085	International filing date (day/month/year) 04/12/1998	(Earliest) Priority Date (day/month/year) 04/12/1997
Applicant INSTITUT PASTEUR et al.		

This International Search Report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This International Search Report consists of a total of 4 sheets.
☒ It is also accompanied by a copy of each prior art document cited in this report.

1. Basis of the report

a. With regard to the **language**, the international search was carried out on the basis of the international application in the language in which it was filed, unless otherwise indicated under this item.

☐ the international search was carried out on the basis of a translation of the international application furnished to this Authority (Rule 23.1(b)).

b. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, the international search was carried out on the basis of the sequence listing :

- ☐ contained in the international application in written form.
- ☐ filed together with the international application in computer readable form.
- ☐ furnished subsequently to this Authority in written form.
- ☐ furnished subsequently to this Authority in computer readable form.
- ☐ the statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.
- ☐ the statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished

2. ☐ **Certain claims were found unsearchable** (See Box I).

3. ☐ **Unity of invention is lacking** (see Box II).

4. With regard to the **title**,

- ☒ the text is approved as submitted by the applicant.
- ☐ the text has been established by this Authority to read as follows:

5. With regard to the **abstract**,

- ☒ the text is approved as submitted by the applicant.
- ☐ the text has been established, according to Rule 38.2(b), by this Authority as it appears in Box II. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority.

6. The figure of the **drawings** to be published with the abstract is Figure No.

- ☐ as suggested by the applicant.
- ☐ because the applicant failed to suggest a figure.
- ☒ because this figure better characterizes the invention.

1c
☐ None of the figures.

INTERNATIONAL SEARCH REPORT

International Application No
IB 98/02085

A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 G01N33/50 C12Q1/02 C12Q1/527 C12Q1/68 G01N33/542
G01N33/58

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 G01N C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>PELLETIER J N ET AL: "A PROTEIN COMPLEMENTATION ASSAY FOR DETECTION OF PROTEIN-PROTEIN INTERACTIONS IN VIVO" PROTEIN ENGINEERING, vol. 10, no. SUPPL. 01, 1 October 1997, page 89 XP002064563 see the whole document</p> <p style="text-align: center;">--- -/--</p>	<p>1,7, 10-17, 20-22, 25-33, 36-38, 40-45</p>

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance
 "E" earlier document but published on or after the international filing date
 "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
 "O" document referring to an oral disclosure, use, exhibition or other means
 "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
 "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
 "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
 "&" document member of the same patent family

Date of the actual completion of the international search

21 April 1999

Date of mailing of the international search report

06/05/1999

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
 NL - 2280 HV Rijswijk
 Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
 Fax: (+31-70) 340-3016

Authorized officer

Gundlach, B

INTERNATIONAL SEARCH REPORT

International Application No

PCT/IB 98/02085

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
----------	--	-----------------------

X	<p>ROSSI F ET AL: "MONITORING PROTEIN-PROTEIN INTERACTIONS IN INTACT EUKARYOTIC CELLS BY BETA-GALACTOSIDASE COMPLEMENTATION" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 94, August 1997, pages 8405-8410, XP002064565 cited in the application see abstract see page 8405, column 2, paragraph 2 - page 8408, column 1, paragraph 2</p>	<p>1,7, 10-17, 20-22, 25-33, 36-38, 40-45</p>
---	---	--

A	<p>LADANT, DANIEL ET AL.: "Characterization of the Calmodulin-binding and of the Catalytic Domains of Bordetella pertussis Adenylate Cyclase" THE JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 264, no. 7, 5 March 1989, pages 4015-4020, XP002100622 cited in the application see abstract see page 4017, column 2, paragraph 3 - page 4019, column 1, paragraph 1 see page 4019, column 1, paragraph 3 - column 2</p>	<p>1-6,8,9, 18,23, 24,34, 35,39</p>
---	--	--

A	<p>GLASER P ET AL: "IDENTIFICATION OF RESIDUES ESSENTIAL FOR CATALYSIS AND BINDING OF CALMODULIN IN BORDETELLA PERTUSSIS ADENYLATE CYCLASE BY SITE-DIRECTED MUTAGENESIS" EMBO JOURNAL, vol. 8, no. 3, 1 March 1989, pages 967-972, XP000082630 see abstract</p>	<p>1-6,8,9, 18,23, 24,34, 35,39</p>
---	--	--

P,X	<p>WO 98 34120 A (PELLETIER JOELLE NINA ;REMY INGRID (CA); UNIV MONTREAL (CA); MICHN) 6 August 1998</p>	<p>1,7, 10-17, 20-22, 25-33, 36-38, 40-45</p>
-----	---	--

A	<p>see abstract</p> <p>see page 5, line 16 - page 16, line 27; figure 1 see page 42, line 4 - page 48, line 8 see page 49, line 16 - page 55, line 27 see claims 1-58; table 1</p>	<p>2-6,8,9, 18,23, 24,34, 35,39</p>
---	--	--

-/--

INTERNATIONAL SEARCH REPORT

International Application No

IB 98/02085

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P, X.	<p>KARIMOVA, GOUZEL ET AL: "A bacterial two-hybrid system based on a reconstituted signal transduction pathway" PROC. NATL. ACAD. SCI. U. S. A. (1998), 95(10), 5752-5756 CODEN: PNASA6; ISSN: 0027-8424, 1998, XP002100623 see the whole document</p>	1-45

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/IB 98/02085

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9834120 A	06-08-1998	CA 2196496 A AU 5850598 A	31-07-1998 25-08-1998

INTERNATIONAL SEARCH REPORT

International Application No
PCT/IB 98/02085

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 G01N33/50 C12Q1/02 C12Q1/527 C12Q1/68 G01N33/542
G01N33/58

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 G01N C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

RECEIVED
AUG 31 2000
TECH CENTER 1600/2900

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>PELLETIER J N ET AL: "A PROTEIN COMPLEMENTATION ASSAY FOR DETECTION OF PROTEIN-PROTEIN INTERACTIONS IN VIVO" PROTEIN ENGINEERING, vol. 10, no. SUPPL. 01, 1 October 1997, page 89 XP002064563 see the whole document</p> <p>---</p> <p>-/--</p>	<p>1,7, 10-17, 20-22, 25-33, 36-38, 40-45</p>

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

21 April 1999

Date of mailing of the international search report

06/05/1999

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Gundlach, B

INTERNATIONAL SEARCH REPORT

International Application No
PCT/IB 98/02085

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
------------	--	-----------------------

X	<p>ROSSI F ET AL: "MONITORING PROTEIN-PROTEIN INTERACTIONS IN INTACT EUKARYOTIC CELLS BY BETA-GALACTOSIDASE COMPLEMENTATION" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 94, August 1997, pages 8405-8410, XP002064565 cited in the application see abstract see page 8405, column 2, paragraph 2 - page 8408, column 1, paragraph 2</p>	<p>1,7, 10-17, 20-22, 25-33, 36-38, 40-45</p>
---	---	---

A	<p>LADANT, DANIEL ET AL.: "Characterization of the Calmodulin-binding and of the Catalytic Domains of Bordetella pertussis Adenylate Cyclase" THE JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 264, no. 7, 5 March 1989, pages 4015-4020, XP002100622 cited in the application see abstract see page 4017, column 2, paragraph 3 - page 4019, column 1, paragraph 1 see page 4019, column 1, paragraph 3 - column 2</p>	<p>1-6,8,9, 18,23, 24,34, 35,39</p>
---	--	---

A	<p>GLASER P ET AL: "IDENTIFICATION OF RESIDUES ESSENTIAL FOR CATALYSIS AND BINDING OF CALMODULIN IN BORDETELLA PERTUSSIS ADENYLATE CYCLASE BY SITE-DIRECTED MUTAGENESIS" EMBO JOURNAL, vol. 8, no. 3, 1 March 1989, pages 967-972, XP000082630 see abstract</p>	<p>1-6,8,9, 18,23, 24,34, 35,39</p>
---	---	---

P,X	<p>WO 98 34120 A (PELLETIER JOELLE NINA ;REMY INGRID (CA); UNIV MONTREAL (CA); MICHN) 6 August 1998</p>	<p>1,7, 10-17, 20-22, 25-33, 36-38, 40-45</p>
-----	---	---

A	<p>see abstract</p> <p>see page 5, line 16 - page 16, line 27; figure 1 see page 42, line 4 - page 48, line 8 see page 49, line 16 - page 55, line 27 see claims 1-58; table 1</p>	<p>40-45 2-6,8,9, 18,23, 24,34, 35,39</p>
---	--	---

-/--

INTERNATIONAL SEARCH REPORT

International Application No
PCT/IB 98/02085

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P, X	<p>KARIMOVA, GOUZEL ET AL: "A bacterial two-hybrid system based on a reconstituted signal transduction pathway" PROC. NATL. ACAD. SCI. U. S. A. (1998), 95(10), 5752-5756 CODEN: PNASA6; ISSN: 0027-8424, 1998, XP002100623 see the whole document -----</p>	1-45

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/IB 98/02085

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9834120 A	06-08-1998	CA 2196496 A AU 5850598 A	31-07-1998 25-08-1998

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



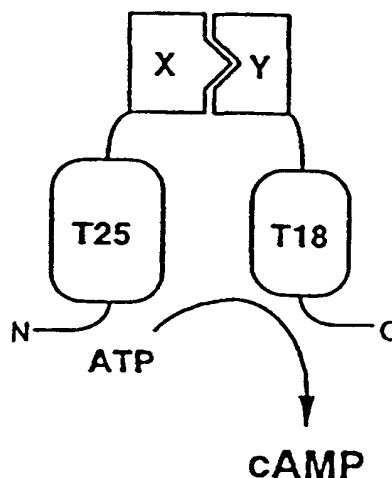
INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : G01N 33/50, C12Q 1/02, 1/527, 1/68, G01N 33/542, 33/58		A1	(11) International Publication Number: WO 99/28746
			(43) International Publication Date: 10 June 1999 (10.06.99)
(21) International Application Number: PCT/IB98/02085		(81) Designated States: CA, JP, US, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).	
(22) International Filing Date: 4 December 1998 (04.12.98)			
(30) Priority Data: 60/067,308 4 December 1997 (04.12.97) US 09/203,681 1 December 1998 (01.12.98) US		Published With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments. With an indication in relation to deposited biological material furnished under Rule 13bis separately from the description.	
(71) Applicant (for all designated States except US): INSTITUT PASTEUR [FR/FR]; 28, rue du Docteur-Roux, F-75015 Paris (FR).			
(72) Inventors; and			
(75) Inventors/Applicants (for US only): LADANT, Daniel [FR/FR]; 3, chemin de Reims, F-94230 Cachan (FR). KARIMOVA, Gouzel [RU/FR]; 11, rue Dupin, F-75006 Paris (FR). ULLMANN, Agnès [FR/FR]; 3, rue Paul-Dupuy, F-75016 Paris (FR).			
(74) Agents: MARTIN, Jean-Jacques et al.; Cabinet Regimbeau, 26, avenue Kléber, F-75116 Paris (FR).			

(54) Title: A BACTERIAL MULTI-HYBRID SYSTEM AND APPLICATIONS

(57) Abstract

A signal amplification system comprises a bacterial multi-hybrid system, and more preferably a two-hybrid system, of at least two chimeric polypeptides containing a first chimeric polypeptide corresponding to a first fragment of an enzyme and a second chimeric polypeptide corresponding to a second fragment of an enzyme or a modulating substance capable of activating said enzyme. The first fragment is fused to a molecule of interest and the second fragment or the modulating substance is fused to a target ligand. The activity of the enzyme is restored by the *in vivo* interaction between the molecule of interest and the target ligand. Signal amplification is generated and, for example, triggers transcriptional activation. The signal amplification system is useful in a method of selecting a molecule of interest, which is capable of binding to target ligand, wherein the interaction between the molecule of interest and the target ligand is detected with the signal amplification system as a kit therefor. A method of screening for a substance capable of stimulating or inhibiting the interaction between a target ligand and a molecule of interest is also provided.



INTERNATIONAL SEARCH REPORT

International Application No
PCT/IB 98/02085

A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 G01N33/50 C12Q1/02 C12Q1/527 C12Q1/68 G01N33/542
G01N33/58

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 G01N C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>PELLETIER J N ET AL: "A PROTEIN COMPLEMENTATION ASSAY FOR DETECTION OF PROTEIN-PROTEIN INTERACTIONS IN VIVO" PROTEIN ENGINEERING, vol. 10, no. SUPPL. 01, 1 October 1997, page 89 XP002064563 see the whole document</p> <p>---</p> <p>-/--</p>	<p>1,7, 10-17, 20-22, 25-33, 36-38, 40-45</p>

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

21 April 1999

Date of mailing of the international search report

06/05/1999

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Gundlach, B

INTERNATIONAL SEARCH REPORT

International Application No

PCT/IB 98/02085

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	ROSSI F ET AL: "MONITORING PROTEIN-PROTEIN INTERACTIONS IN INTACT EUKARYOTIC CELLS BY BETA-GALACTOSIDASE COMPLEMENTATION" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 94, August 1997, pages 8405-8410, XP002064565 cited in the application see abstract see page 8405, column 2, paragraph 2 - page 8408, column 1, paragraph 2 ---	1,7, 10-17, 20-22, 25-33, 36-38, 40-45
A	LADANT, DANIEL ET AL.: "Characterization of the Calmodulin-binding and of the Catalytic Domains of Bordetella pertussis Adenylate Cyclase" THE JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 264, no. 7, 5 March 1989, pages 4015-4020, XP002100622 cited in the application see abstract see page 4017, column 2, paragraph 3 - page 4019, column 1, paragraph 1 see page 4019, column 1, paragraph 3 - column 2 ---	1-6,8,9, 18,23, 24,34, 35,39
A	GLASER P ET AL: "IDENTIFICATION OF RESIDUES ESSENTIAL FOR CATALYSIS AND BINDING OF CALMODULIN IN BORDETELLA PERTUSSIS ADENYLATE CYCLASE BY SITE-DIRECTED MUTAGENESIS" EMBO JOURNAL, vol. 8, no. 3, 1 March 1989, pages 967-972, XP000082630 see abstract ---	1-6,8,9, 18,23, 24,34, 35,39
P,X	WO 98 34120 A (PELLETIER JOELLE NINA ;REMY INGRID (CA); UNIV MONTREAL (CA); MICHN) 6 August 1998 ---	1,7, 10-17, 20-22, 25-33, 36-38, 40-45
A	see abstract see page 5, line 16 - page 16, line 27; figure 1 see page 42, line 4 - page 48, line 8 see page 49, line 16 - page 55, line 27 see claims 1-58; table 1 ---	2-6,8,9, 18,23, 24,34, 35,39
	--- -/--	



.

J

.

.

7

2

INTERNATIONAL SEARCH REPORT

International Application No
PCT/IB 98/02085

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	<p>KARIMOVA, GOUZEL ET AL: "A bacterial two-hybrid system based on a reconstituted signal transduction pathway"</p> <p>PROC. NATL. ACAD. SCI. U. S. A. (1998), 95(10), 5752-5756 CODEN: PNASA6; ISSN: 0027-8424, 1998, XP002100623</p> <p>see the whole document</p> <p>-----</p>	1-45

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/IB 98/02085

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 9834120	A	06-08-1998	CA 2196496 A	31-07-1998
			AU 5850598 A	25-08-1998

7pts

WO 99/28746

09/555649
Rec'd PCT/IB98/02085 02 JUN 2000

1

A BACTERIAL MULTI-HYBRID SYSTEM AND APPLICATIONS

BACKGROUND OF THE INVENTION

5 The present invention concerns a method for selecting a molecule, and kit thereof, a method for screening a molecule, and kit thereof, and a signal amplification system comprising a bacterial multi-hybrid system.

10 The present invention relates to a signal amplification system comprising a bacterial multi-hybrid system, and more preferably a two-hybrid system, of at least two chimeric polypeptides containing a first chimeric polypeptide corresponding to a first fragment of an enzyme and a second chimeric polypeptide corresponding to a second fragment of an enzyme or a modulating substance capable of activating said
15 enzyme, wherein the first fragment is fused to a molecule of interest and the second fragment or the modulating substance is fused to a target ligand, and wherein the activity of the enzyme is restored by the interaction between the said molecule of interest and the said target ligand, and wherein
20 a signal amplification is generated.

The present invention also relates to a method of selecting a molecule of interest, which is capable of binding to a target ligand, wherein the interaction between the said molecule of interest and the said target ligand is detected
25 with a signal amplification system according to the invention, by means of generating a signal amplification and triggering transcriptional activation.

The present invention also relates to a method of screening for a substance capable of stimulating or
30 inhibiting the interaction between a target ligand and a molecule of interest, wherein respectively the stimulating or the inhibiting activity is detected with a signal amplification system according to the invention, by means of generating a signal amplification and respectively of
35 triggering or of abolishing transcriptional activation, and



wherein said signal amplification and said triggered or abolished transcriptional activation are compared with those obtained from an identical signal amplification system without any substance.

5 Most biological processes involve specific protein-protein interactions. General methodologies to identify interacting proteins or to study these interactions have been extensively developed. Among them, the yeast two-hybrid system currently represents the most powerful *in vivo*
10 approach to screen for polypeptides that could bind to a given target protein. Originally developed by Fields and coworkers [Fields, S. & Song, O. (1989) *Nature* 340, 245-6; Chien, C. T., Bartel, P. L., Sternglanz, R. & Fields, S. (1991) *Proc. Natl. Acad. Sci. USA* 88, 9578-82. Two American
15 Patents n° 5,283,173 granted on February 1, 1994 (Fields, S. & Song, O.) and n° 5,468,614 granted on November 21, 1995 (Fields, S. & Song, O.) are also incorporated by reference], it utilizes hybrid genes to detect protein-protein interactions by means of direct activation of a reporter-gene
20 expression (Allen, J. B., Walberg, M. W., Edwards, M. C. & Elledge, S. J. (1995) *Trends Biochem. Sci.* 20, 511-6; Transy, C. & Legrain, P. (1995) *Mol. Biol. Rep.* 21, 119-27).

In essence, the two putative protein partners are genetically fused to the DNA-binding domain of a
25 transcription factor and to a transcriptional activation domain, respectively. A productive interaction between the two proteins of interest will bring the transcriptional activation domain in the proximity of the DNA-binding domain and will trigger directly the transcription of an adjacent
30 reporter gene (usually *lacZ* or a nutritional marker) giving a screenable phenotype. As there is evidence that the transcription can be activated through the use of two functional domains of a transcription factor: a domain that recognizes and binds to a specific site on the DNA and a

domain that is necessary for activation, as reported by Keegan et al. (1986) *Science* 231, 699-407 and Ma and Ptashne (1987) *Cell* 48, 847-853.

Recently, Rossi et al. (Rossi, F., Charlton, C. A. & Blau, H. M. (1997) *Proc. Natl. Acad. Sci. USA.* 94, 8405-8410) described a different approach, a mammalian "two-hybrid" system, which uses β -galactosidase complementation (Ullmann, A., Jacob, F. & Monod, J. (1968) *J. Mol. Biol.* 32, 1-13) to monitor protein-protein interactions in intact eukaryotic cells.

Phage display (Smith, G. P. (1985) *Science* 228, 1315-7; Scott, J. K. & Smith, G. P. (1990) *Science* 249, 386-90) and double-tagging assay (Germino, F. J., Wang, Z. X. & Weissman, S. M. (1993) *Proc. Natl. Acad. Sci. USA.* 90, 933-7) represent alternative approaches to screen complex libraries of proteins for direct interaction with a given ligand. However, these techniques do not allow an *in vivo* selection of the relevant clones.

Another approach is described in the International Patent Application n° WO 96/40987 (Schatz, P. J. et al.), which provides random peptide libraries and methods for generating and screening libraries to identify peptides that bind to receptor molecules of interest, including antibodies. The peptide library is constructed so that the DNA binding protein-random peptide fusion product can bind to the recombinant DNA expression vector that encodes the fusion product that contains the peptide of interest. The method of generating the peptide library comprises the steps of (a) constructing a recombinant DNA vector that encodes a DNA binding protein and contains binding sites for the DNA binding protein; (b) inserting into the coding sequence of the DNA binding protein in a multiplicity of vectors of step (a) coding sequences for random peptides such that the resulting vectors encode different fusion proteins, each of



which is composed of the DNA binding protein and a random peptide; (c) transforming host cells with the vectors of steps (b); and (d) culturing the host cells transformed in step (c) under conditions suitable for expression of the fusion proteins. Typically, a random peptide library will contain at least 10^6 to 10^8 different members, although library sizes of 10^8 to 10^{11} can be achieved.

A novel variety of approach is defined in the International Patent Application n° WO 96/29429 (Wickens, M. & Fields, S.) related to a hybrid system to detect protein-RNA interactions using the same method of achievement as recited in the two above-mentioned American patents. This hybrid system has a first hybrid protein comprising a DNA-binding domain and a first RNA-binding domain, a second hybrid protein comprising a transcriptional activation domain and a second RNA-binding domain, and a hybrid RNA. The interaction between both the first RNA-binding domain and the hybrid RNA and the second RNA-binding domain and the hybrid RNA causes the transcriptional activation domain to activate transcription of the detectable gene.

Bartel, P. L., Roecklein, J. A., SenGupta, D. & Fields, S. (1996) *Nat. Genet.* 12, 72-77 extended the approach of the typical two-hybrid system consisting in a known protein that forms a part of a DNA-binding domain hybrid, assayed against a library of all possible proteins present as transcriptional activation domain hybrids, using the genome of the bacteriophage T7, such that a second library of all possible proteins is fused to the DNA-binding domain to be analyzed. This genome-wide approach to the two-hybrid searches has identified 25 interactions among the proteins of T7.

SUMMARY OF THE INVENTION

The aim of the present invention is to provide a novel bacterial multi-hybrid system, and more preferably a two-



hybrid system, in which proteins of interest are genetically fused to two complementary fragments of a catalytic domain of an enzyme, which provides significant advantages over the prior art.

5 Thus, the present invention provides a signal amplification system comprising a bacterial multi-hybrid system, and more preferably a two-hybrid system, of at least two chimeric polypeptides containing a first chimeric polypeptide corresponding to a first fragment of an enzyme
10 and a second chimeric polypeptide corresponding to a second fragment of an enzyme or a modulating substance capable of activating said enzyme, wherein the first fragment is fused to a molecule of interest and the second fragment or the modulating substance is fused to a target ligand, and wherein
15 the activity of the enzyme is restored by the interaction between the said molecule of interest and the said target ligand, and wherein a signal amplification is generated.

 This system allows an easy *in vivo* screening and selection of functional interactions between the target
20 ligand and the molecule of interest.

 A genetic test is based on the reconstitution, in a specific enzyme deficient bacteria, of a signal transduction pathway that takes advantage of the positive control exerted by a signaling molecule. Association of the target ligand and
25 the molecule of interest results in functional complementation between the two chimeric polypeptides and leads to the signaling molecule synthesis. The signaling molecule then triggers transcriptional activation of catabolic operons, of a gene conferring resistance to
30 antibiotics, of a gene encoding for a toxin or of a color marker, such as a fluorescent marker of the type of the Green Fluorescent Protein (GFP) that yields a characteristic phenotype. In this genetic test of screening and/or selection, the involvement of a signaling cascade offers the
35 unique property that association between the chimeric



polypeptides can be spatially separated from the transcriptional activation readout. This permits a versatile design of screening procedures either for ligands that bind to a given "bait", as in the classical yeast multi-hybrid system, or for molecules or mutations that block a given interaction between two proteins of interest.

Furthermore, because the signal amplification system according to the invention involves the generation of at least one signaling molecule, also called regulatory molecule, the physical association of the two putative interacting target ligand and molecule of interest can be spatially separated from the transcriptional events that are dependent on regulatory molecule synthesis. This means that the interaction between a target ligand and a molecule of interest under study does not need to take place in the vicinity of the transcription machinery as is the case for the yeast two-hybrid system as described above. Hence, in addition to the methods described above, the present invention allows one to analyze more particularly protein interactions that occur either in the cytosol or at the inner membrane level.

Another advantage of the present invention over the prior art is that this bacterial system is particularly versatile as it offers the possibility of both positive and negative selections. Positive selection means bacterial growth, for example, on minimal medium containing lactose or maltose. Negative selection means arrest of growth.

The present invention further relates to a method of selecting a molecule of interest, which is capable of binding to a target ligand, wherein the interaction between the said molecule of interest and the said target ligand is detected with a signal amplification system according to the invention, by means of generating a signal amplification and triggering transcriptional activation.



The present invention also relates to a method of screening for a substance capable of stimulating or inhibiting the interaction between a target ligand and a molecule of interest, wherein the stimulating or the inhibiting activity is detected with a signal amplification system according to the invention, by means of generating a signal amplification and triggering transcriptional activation, and wherein said signal amplification and said triggering transcriptional activation are compared with those obtained from an identical signal amplification system without any substance.

The present invention also provides a kit for selecting a molecule of interest, wherein said kit comprises:

(a) a signal amplification system according to the invention;

(b) an *E. coli* strain, or a bacterial strain, or an eukaryotic cell deficient in endogenous adenylate cyclase; and

(c) a medium allowing the detection of the complementation selected from the group consisting of indicator or selective medium, for example, as minimal medium supplemented with lactose or maltose as unique carbon source, medium with antibiotics, medium to visualize fluorescence, conventional medium, and medium that allows the sorting by the presence of the phage receptor.

Further, the present invention also provides a kit for selecting a molecule of interest, wherein said kit comprises:

(a) a signal amplification system according to the invention, wherein the molecule of interest is a mutant molecule compared to the known wild type molecule;

(b) a signal amplification system according to the invention, wherein the molecule of interest is the known wild type molecule as the control;



(c) *E. coli* strain, or in any bacterial strain deficient in endogenous adenylate cyclase, or any other eukaryotic cell;

5 (d) a medium allowing the detection of the complementation selected from the group consisting of indicator plate or selective medium as minimal medium supplemented with lactose or maltose as unique carbon source medium with antibiotics, medium to visualize fluorescence, conventional medium, and medium that allows the sorting by
10 the presence of the phage receptor for each signal amplification system; and

(e) means for detecting whether the signal amplification system with the mutant molecule is enhanced or inhibited with respect to the signal amplification system with the wild type
15 molecule.

The present invention also provides a kit for screening for a substance capable of stimulating or inhibiting the interaction between a target ligand and a molecule of interest, wherein said kit comprises:

20 (a) a signal amplification system according to the invention with the substance capable of stimulating or inhibiting the interaction between a target ligand and a molecule of interest;

(b) a signal amplification system according to the
25 invention without any substance as the control;

(c) *E. coli* strain, or in any bacterial strain deficient in endogenous adenylate cyclase, or any other eukaryotic cell and;

(d) a medium allowing the detection of the
30 complementation selected from the group consisting of indicator plate or selective medium as minimal medium supplemented with lactose or maltose as unique carbon source, medium with antibiotics, medium to visualize fluorescence, conventional medium, and medium which allows the sorting by
35 the presence of the phage receptor;



(e) means for detecting whether the signal amplification system with the substance is enhanced or inhibited with respect to the signal amplification system without any substance.

5 According to one embodiment of the present invention, the signal amplification system comprises a bacterial multi-hybrid system, and more preferably a two-hybrid system, containing a first chimeric polypeptide corresponding to a first fragment of an enzyme, a second chimeric polypeptide
10 corresponding to a second fragment of an enzyme or a modulating substance capable of activating said enzyme, and a substance capable of stimulating or inhibiting the interaction between a target ligand and a molecule of interest.

15 The present invention also provides a method of screening for a substance capable of stimulating or inhibiting the interaction between a target ligand and a molecule of interest, wherein the stimulating or the inhibiting activity is detected with a signal amplification
20 system according to the invention, by means of generating a signal amplification and triggering transcriptional activation, and wherein said signal amplification and said triggering transcriptional activation are compared with those obtained from an identical signal amplification system
25 without any substance.

BRIEF DESCRIPTION OF THE DRAWINGS

This invention will be described in greater detail with reference to the drawings in which:

30 Figures 1A, 1B, 1C and 1D depict the principle of an *E. coli* multi-hybrid system based on functional complementation of the catalytic domain of *Bordetella* adenylate cyclase (CyaA) fragments.



The upper part schematizes the basic principle of *in vivo* complementation between the two fragments of the catalytic domain of *B. pertussis* adenylate cyclase. The two boxes represent the T25 and T18 fragments corresponding to amino acids 1 to 224 and 225 to 399 of the CyaA protein. In figure 1A, the full-length catalytic domain (residues 1 to 399), when expressed in *E. coli*, exhibits a basal calmodulin-independent activity that results in cyclic adenosine monophosphate (cAMP) synthesis. In figure 1B, the two fragments T25 and T18, when coexpressed as independent polypeptides, are unable to interact and no cAMP synthesis occurs. In figure 1C, the two fragments, fused to two interacting proteins, X and Y, are brought into close proximity resulting in functional complementation, followed by cAMP production.

The lower part schematizes the readout of the complementation. cAMP, synthesized in an *E. coli* *cya* strain by the complementing T25 and T18 pairs, binds to the catabolite activator protein, CAP. The cAMP/CAP complex (C) can then recognize specific promoters and switch on the transcription of the corresponding genes. These reporter genes can be either natural *E. coli* genes, such as *lacZ* or *mal* genes, or synthetic ones, such as antibiotic resistance genes fused to a cAMP/CAP dependent promoter.

Figure 2 is a schematic representation of plasmids.

The open boxes represent the open reading frames of β -lactamase (*bla*) and chloramphenicol acetyl transferase (*cat*) genes. The dark boxes correspond to the open reading frame of *cyaA'* with codon numbers indicated below. The hatched boxes correspond to the multicloning site sequences (MCS) that are fused at the indicated position of the *cya* open reading frame. The origin of replication of the plasmids is indicated by dotted boxes.



Figure 3.1 and Figure 3.2 are schematic representations of other plasmids.

The left part represents the maps of the plasmids, with the different antibiotic-selectable markers (chloramphenicol acetyl transferase (*cat*), aminoglycoside phosphotransferase (*kan*) and β -lactamase (*bla*), the origin of replication and the position of the multicloning site sequences (MCS) relative to the T25 and T18 open reading frames. The right part describes the nucleotide sequence of the multicloning site sequences (MCS) fused to T25 (Fig. 3.2) or T18 (Fig. 3.1) and the corresponding reading frames.

Figures 4A and 4B depict the results of screening of interacting proteins with the bacterial two-hybrid system.

DHPI cells were cotransformed with a mixture of plasmids pT18, pT18-*zip*, and pT18-Tyr, and either pT25 (Fig. 4A) or pT25-*zip* (Fig. 4B), plated on LB-X-Gal agar plates containing 0.5 mM IPTG, ampicillin and chloramphenicol, and incubated for 30 hrs at 30°C. Note that the *cya*⁺ colonies are larger than the *cya*⁻ ones.

Figure 5 relates to the mapping of interacting domains of the *B. stearrowthermophilus* tyrosyl-tRNA synthetase.

DNA fragments encoding the indicated polypeptide segments of the tyrosyl-tRNA synthetase (the numbers correspond to the amino acid residues) were amplified by PCR using appropriate primers and cloned into pT25 and/or pT18. The functional complementation between the indicated chimeric proteins was assayed on DHPI cells co-transformed with the corresponding plasmids by measuring the β -galactosidase activity.

Figure 6 relates to the mapping of interacting domains of *B. pertussis* B.vgA.

DNA fragments encoding indicated polypeptide segments of BvgA (the numbers correspond to the amino acid residues) were amplified by PCR using appropriate primers and cloned into



pKT25 and/or pUTT18C. The functional complementation between the indicated chimeric proteins was assayed on DHP1 cells co-transformed with the corresponding plasmids by measuring the β -galactosidase activity.

5

DETAILED DESCRIPTION OF THE INVENTION

Thus, the present invention provides a novel signal amplification system in *Escherichia coli*, in which the proteins of interest are genetically fused to two complementary fragments of the catalytic domain of *Bordetella pertussis* adenylate cyclase (Ladant, D. (1988) *J. Biol. Chem.* 263, 2612-2618; Ladant, D., Michelson, S., Sarfati, R. S., Gilles, A.-M., Predeleanu, R. & Blrzu, O. (1989) *J. Biol. Chem.* 264, 4015-4020).

15 *B. pertussis* produces a calmodulin dependent adenylate cyclase toxin encoded by the *cyaA* gene (Hewlett, E. L., Urban, M. A., Manclark, C. R. & Wolff, J. (1976) *Proc. Natl. Acad. Sci. U.S.A.* 73, 1926-1930; Glaser, P., Ladant, D., Sezer, O., Pichot, F., Ullmann, A. & Danchin, A. (1988) *Mol. Microbiol.* 2, 19-30; Mock, M. & Ullmann, A. (1993) *Trends Microbiol.* 1, 187-192). The catalytic domain is located within the first 400 amino acids of this 1706 residue-long protein (Ladant, D., Michelson, S., Sarfati, R. S., Gilles, A.-M., Predeleanu, R. & Blrzu, O. (1989) *J. Biol. Chem.* 264, 4015-4020; Glaser, P., Ladant, D., Sezer, O., Pichot, F., Ullmann, A. & Danchin, A. (1988) *Mol. Microbiol.* 2, 19-30). It exhibits a high catalytic activity ($k_{cat} = 2000 \text{ s}^{-1}$) in the presence of calmodulin (CaM), and a low but detectable activity ($k_{cat} = 2 \text{ s}^{-1}$) in the absence of this activator (Ladant, D. (1988) *J. Biol. Chem.* 263, 2612-2618; Wolff, J., Cook, G. H., Goldhammer, A. R. & Berkowitz, S. A. (1980) *Proc. Natl. Acad. Sci. USA.* 77, 3841-3844).

30

Biochemical studies revealed that the catalytic domain can be proteolytically cleaved into two complementary



fragments, T25 and T18, that remain associated in the presence of CaM in a fully active ternary complex (Ladant, D. (1988) *J. Biol. Chem.* 263, 2612-2618; Ladant, D., Michelson, S., Sarfati, R. S., Gilles, A.-M., Predeleanu, R. & Blrzu, O. (1989) *J. Biol. Chem.* 264, 4015-4020; Munier, H., Gilles, A. M., Glaser, P., Krin, E., Danchin, A., Sarfati, R. & Barzu, O. (1991) *Eur. J. Biochem.* 196, 469-74). In the absence of CaM, the mixture of the two fragments did not exhibit detectable activity suggesting that the two fragments are not able to reassociate to yield basal CaM-independent activity.

The two complementary fragments, T25 and T18, that are both necessary to form an active enzyme, in the presence of CaM when expressed in *E. coli* as separated entities, are unable to recognize each other and cannot reconstitute a functional enzyme. However, when T25 and T18 are fused to peptides or proteins that are able to interact, heterodimerization of these chimeric polypeptides results in a functional complementation between the adenylate cyclase fragments.

When expressed in an adenylate cyclase deficient *E. coli* strain (*E. coli* lacks CaM or CaM-related proteins), the T25 and T18 fragments fused to putative interacting proteins reassociate and lead to cAMP synthesis (Figures 1A, 1B and 1C).

Interaction between a target ligand and a molecule of interest results in functional complementation between the two adenylate cyclase fragments leading to cAMP synthesis, which in turn can trigger the expression of several resident genes. Using this assay, one can select specific clones expressing a protein that interacts with a given target by a simple genetic screening.

The present invention provides a signal amplification system comprising a bacterial multi-hybrid system, and more preferably a two-hybrid system, of at least two chimeric



polypeptides containing a first chimeric polypeptide corresponding to a first fragment of an enzyme, and a second chimeric polypeptide corresponding to a second fragment of an enzyme or a modulating substance capable of activating said enzyme, wherein the first fragment is fused to a molecule of interest and the second fragment or the modulating substance is fused to a target ligand, and wherein the activity of the enzyme is restored by the interaction between the said molecule of interest and the said target ligand, and wherein a signal amplification is generated.

"Signal amplification system" means a system involving the interaction between at least two chimeric polypeptides leading to the production of a large number of signaling molecules.

"Signal amplification" means, in the present invention, that the number of signaling molecules is higher than the number of chimeric polypeptides that produced it.

The first fragment and the second fragment are issued from the same enzyme or not. In any case, the first and the second fragments are distinct from each other even if they are issued from the same enzyme. For example, the fragments are from residues 1 to 224 and 225 to 399 from *B. Pertussis* adenylate cyclase.

A fragment issued from the enzyme comprises between 20 and 400 amino acid residues and more preferably 200 consecutive amino acid residues.

"Modulating substance" refers to a substance capable of activating or inhibiting an enzyme, which is an activator, natural or not, of the enzyme a fragment thereof, or a derivative of the activator; the enzyme having a modulating substance-binding site. In a preferred embodiment of the invention, the modulating substance is a natural activator as, for example, the calmodulin.

The fragments and the modulating substance are fused, respectively, to the molecule of interest or to the target



ligand by means of genetic recombination as described herein after. A proteolytic cleavage site can be introduced, according to the well known techniques, in the genetic construction between a fragment of the enzyme and a molecule
5 of interest in order to eliminate easily, after the generation of the signal amplification, by restriction enzyme the fragment and to recover the molecule of interest.

The molecule of interest can be detected for example from cDNA, genomic, or synthetic random DNA libraries.

10 The restoration of the enzymatic activity means that an enzyme activity is recovered.

The interaction between the molecule of interest and the target ligand means that there exists a recognition which could possibly lead to the binding between the molecule of
15 interest and the target ligand.

According to the invention, the enzyme is selected from the group consisting of adenylate cyclase and guanylate cyclase from any origin. Any origin refers to *Bordetella* species or any other organism that produces this type of
20 enzyme.

In one specific illustration of the present invention the enzyme is the catalytic domain of *Bordetella* adenylate cyclase (CyaA) located within the first 400 amino acid residues of the adenylate cyclase toxin.

25 The present invention also concerns a first fragment and a second fragment, which are any combination of fragments from the same enzyme, which *in vitro* functionally interact with the natural activator of said enzyme by restoring its activity.

30 According to one embodiment of the invention the first and the second fragments are selected from the group consisting of :

(a) a fragment T25 corresponding to amino acids 1 to 224 of CyaA and a fragment T18 corresponding to amino acids 225
35 to 399 of CyaA;



(b) a fragment corresponding to amino acids 1 to 224 of CyaA and a fragment corresponding to amino acids 224 to 384 of CyaA;

5 (c) a fragment corresponding to amino acids 1 to 137 of CyaA and a fragment corresponding to amino acids 138 to 400 of CyaA;

(d) a fragment corresponding to amino acids 1 to 317 of CyaA and a fragment corresponding to amino acids 318 to 400 of CyaA; and

10 (e) two fragments from eukaryotic adenylate cyclase in association with molecules, such as, G protein and forskolin.

According to a preferred embodiment of the invention, the first and the second fragments are a fragment T25 corresponding to amino acids 1 to 224 of *Bordetella pertussis* CyaA and a fragment T18 corresponding to amino acids 225 to 399 of *Bordetella pertussis* CyaA.

According to the invention, the modulating substance is a natural activator, or a fragment thereof, of the enzyme. In a specific embodiment of the invention, the natural activator is the calmodulin (CaM), or a fragment thereof, and said first fragment is mutated compared to the wild type enzyme. The fragment of calmodulin is about 70 amino acids long, corresponding preferentially, to residues 77 to 148 of mammalian calmodulin.

25 According to another aspect, the invention also concerns a first fragment, which is a mutated fragment of the catalytic domain of *Bordetella* adenylate cyclase (CyaA). "Mutated fragment" means that it presents at least one mutation in the polynucleotide sequence, said fragment in combination with a second fragment can then *in vivo* functionally interact only in the presence of the natural activator.

30 The signal amplification system according to the invention comprises a bacterial multi-hybrid system containing:



(a) a first chimeric polypeptide corresponding to a first fragment of an enzyme;

(b) a second chimeric polypeptide corresponding to a second fragment of an enzyme or a modulating substance capable of activating said enzyme; and

(c) a substance capable of stimulating or inhibiting the interaction between a target ligand and a molecule of interest, wherein the first fragment is fused to a molecule of interest and the second fragment or the modulating substance is fused to a target ligand, and wherein the activity of the enzyme is restored by the interaction between the said molecule of interest and the said target ligand, and wherein a signal amplification is generated.

Another aspect of the present invention consists in a method of selecting a molecule of interest, which is capable of binding to target ligand, wherein the interaction between the said molecule of interest and the said target ligand is detected with a signal amplification system according to the invention, by means of signal amplification which triggers transcriptional activation, and is quantified by measuring the synthesis of the signaling molecule or the expression of the reporter gene.

This method of selecting a molecule of interest allows selection of a molecule capable of interacting directly with a predetermined target ligand.

The signal amplification corresponds to the production of a signaling molecule. This signaling molecule is any molecule capable of leading to a signaling cascade reaction.

In a preferred embodiment of the invention, the signaling molecule corresponds to the synthesis of cAMP.

In another preferred embodiment of the invention, the signaling molecule corresponds to the synthesis of cGMP.

The transcriptional activation leads to a reporter gene, expression of which is selected from the group consisting of gene coding for nutritional marker, such as lactose or



maltose; gene conferring resistance to antibiotics such as ampicillin, chloramphenicol, kanamycin, or tetracyclin; a gene encoding for a toxin; a color marker, such as, fluorescent marker of the type of the Green Fluorescent Protein (GFP); a gene encoding phage receptor proteins or a
5 fragment thereof, such as phage 8 receptor, *lamB*, and any other gene giving a selectable phenotype.

According to a preferred embodiment of the invention, cAMP, upon binding to CAP, is able to activate the
10 transcription of catabolic operons, allowing the bacteria to ferment carbohydrates, such as maltose or lactose, and to express the phage 8 receptor, protein LamB, which could serve as a marker at the bacterial surface. This signal amplification system comprising this bacterial multi-hybrid
15 system is able to reveal, for example, interactions between small peptides (GCN4 leucine zipper), bacterial (tyrosyl tRNA synthetase), or eukaryotic proteins (yeast Prp11/Prp21 complex).

Accordingly, specific reporter cassettes in which any
20 gene of interest is fused to a cAMP/CAP dependent promoter can be designed. Thus, to facilitate the screening and the selection of complex libraries, the construction of such a simple selection system using an antibiotic resistance gene can be performed.

25 The reporter gene could be a toxin, not naturally present in bacteria, under the control of a cAMP/CAP-dependent promoter. This could be particularly useful to search for chemical compounds or mutations that abolish a given interaction between the target ligand and a molecule of
30 interest. According to this construction, when association between the target ligand and a molecule of interest takes place, cAMP will be produced, the expression of the toxin gene will be switched on, and the cells will be killed. A substance capable of stimulating or inhibiting the
35 interaction between the target ligand and the molecule of



interest and that abolishes interaction will shut down toxin gene expression and will enable the cells to grow. An easy selection for substances that abolish interaction between the target ligand and the molecule of interest is resistance to phage 8. The phage receptor, the LamB protein, is the product of the *lamB* gene, which is part of the maltose regulon, therefore its expression requires cAMP. In consequence, cells producing cAMP will lyse when infected with 8 vir. Substances that abolish interaction between the target ligand and the molecule of interest will abrogate cAMP synthesis and bacteria will become resistant to phage 8. As a result, the cells will grow.

Another selection scheme for compounds or mutations that abolish a given interaction could be designed by constructing a strain that harbors a selectable marker (i.e. a gene conferring resistance to antibiotics such as ampicillin chloramphenicol, kanamycin, tetracyclin, etc.) under the transcriptional control of a promoter that is repressed by cAMP/CAP. Such cAMP/CAP repressed promoter can be engineered by introducing a synthetic CAP binding site within the promoter region as shown by Morita et al. (Morita T, Shigesada K., Kimizuka F., Aiba H. (1988), "Regulatory effect of a synthetic CRP recognition sequence placed downstream of a promoter," Nucleic Acids Res. 16:7315-32).

The International Patent Applications n° WO 96/23898 (Thastrup O. et al.) and n° WO 97/11094 (Thastrup O. et al.), respectively, relating to a method of detecting biologically active substances as Green Fluorescent Protein (GFP), and the International Patent Application n° WO 97/07463 (Chalfie M. et al.) describing the uses of GFP, are herein incorporated by reference, and a novel variant of GFP.

In one specific illustration of the present invention, the method of selecting a molecule of interest consists in a signal amplification system, which comprises a bacterial



multi-hybrid system of at least two distinct fragments of an enzyme, whose enzymatic activity is restored by the interaction between the said molecule of interest and the said target ligand. The two fragments are any combination of
5 fragments from the enzyme, which *in vitro* functionally interact with the natural activator of said enzyme by restoring its activity.

According to the method of selecting a molecule of interest of the present invention, the fragments are selected
10 from the group consisting of :

(a) a fragment T25 corresponding to amino acids 1 to 224 of CyaA and a fragment T18 corresponding to amino acids 225 to 399 of CyaA;

(b) a fragment corresponding to amino acids 1 to 224 of
15 CyaA and a fragment corresponding to amino acids 224 to 384 of CyaA;

(c) a fragment corresponding to amino acids 1 to 137 of CyaA and a fragment corresponding to amino acids 138 to 400 of CyaA;

(d) a fragment corresponding to amino acids 1 to 317 of
20 CyaA and a fragment corresponding to amino acids 318 to 400 of CyaA; and

(e) two fragments from eukaryotic adenylate cyclase in association with molecules, such as G protein and forskolin.

25 And more particularly, the fragments are a fragment T25 corresponding to amino acids 1 to 224 of *Bordetella pertussis* CyaA and a fragment T18 corresponding to amino acids 225 to 399 of *Bordetella pertussis* CyaA.

30 In another specific illustration of the present invention, the method of selecting a molecule of interest consists in a signal amplification system, which comprises a bacterial multi-hybrid system of at least a first fragment of an enzyme and a modulating substance, whose activity, which is an enzymatic activity, is restored by the interaction



between the said molecule of interest and the said target ligand.

In both of the above illustrations of the present invention, the enzyme is selected from the group consisting
5 of adenylate cyclase and guanylate cyclase from any origin, and more preferably the enzyme is the catalytic domain of *Bordetella* adenylate cyclase (CyaA) located within the first 400 amino acid residues of the adenylate cyclase toxin.

The target ligand according to the invention is selected
10 from the group consisting of protein, peptide, polypeptide, receptor, ligand, antigen, antibody, DNA binding protein, glycoprotein, lipoprotein and recombinant protein.

"Peptide" or "polypeptide" or "protein" refers to a polymer in which the monomers are alpha amino acids joined
15 together through amide bonds. Peptides are two or often more amino acid monomers long. Polypeptides are more than ten amino acid residues. Proteins are more than thirty amino acid residues. Standard abbreviations for amino acids are used herein (see Stryer, 1988, *Biochemistry*, Third Ed.,
20 incorporated herein by reference).

"DNA Binding Protein" refers to a protein that specifically interacts with deoxyribonucleotide strands. A sequence specific DNA binding protein binds to a specific
25 sequence or family of specific sequences showing a high degree of sequence identity with each other (e.g., at least about 80 % sequence identity) with at least 100-fold greater affinity than to unrelated sequences. The dissociation constant of a sequence-specific DNA binding protein to its specific sequence(s) is usually less than about 100 nM, and
30 may be as low as 10 nM, 1 nM, 1 pM, or 1 fM. A nonsequence specific DNA binding protein binds to a plurality of unrelated DNA sequences with a dissociation constant that varies by less than 100-fold, usually less than tenfold, to the different sequences. The dissociation constant of a
35 nonsequence specific DNA binding protein to the plurality of



sequences is usually less than about 1:m. In the present invention, DNA binding protein can also refer to an RNA binding protein.

5 "Recombinant protein" refers to a protein made up of at least two separate amino acid chains, which are naturally not contiguous. For example, any fusion protein like Lac repressor- β -galactosidase, any protein or polypeptide like the tyrosyl-tRNA synthetase like leucine zipper derived from protein GCN4.

10 According to the method of selecting a molecule of interest of the present invention, the molecule of interest is capable of interacting with the target ligand and possibly of binding to said target ligand.

15 In a specific embodiment of the method of selecting a molecule of interest of the present invention, the molecule of interest is a mutant molecule compared to the known wild type molecule, and said molecule of interest is tested for its capacity of interacting with the target ligand.

20 The present invention further relates to a kit for selecting a molecule of interest, wherein said kit comprises:

(a) a signal amplification system according to the invention;

25 (b) an *E. coli* strain, or any bacterial strain deficient in endogenous adenylate cyclase, or any other eukaryotic cell; and

30 (c) a medium allowing the detection of the complementation selected from the group consisting of indicator plate or selective medium as minimal medium supplemented with lactose or maltose as unique carbon source, or medium with antibiotics, or medium to visualize fluorescence, conventional medium, and medium which allows sorting by the presence of the phage receptor. The indicator plate is, for example, a MacConkey agar medium supplemented with lactose or maltose.



A bacterial strain deficient in endogenous adenylate cyclase means that this strain is not capable of cAMP synthesis.

The present invention also relates to a kit for
5 selecting a molecule of interest, wherein said kit comprises:

(a) a signal amplification system according to the invention, wherein the molecule of interest is a mutant molecule compared to the known wild type molecule;

(b) a signal amplification system according to the
10 invention, wherein the molecule of interest is the known wild type molecule as the control;

(c) *E. coli* strain, or any bacterial strain deficient in endogenous adenylate cyclase or any other eukaryotic cell; and

(d) a medium allowing the detection of complementation
15 selected from the group consisting of indicator or selective medium as minimal medium supplemented with lactose or maltose as unique carbon source, medium with antibiotics, medium to visualize fluorescence, conventional medium and medium which
20 allows the sorting by the presence of the phage receptor for each signal amplification system; and

(e) means for detecting whether the signal amplification system with the mutant molecule is enhanced or inhibited with respect to the signal amplification system with wild type
25 molecule.

The present invention includes a molecule of interest identified by the method of selecting a molecule of interest according to the present invention.

The present invention further includes a molecule of
30 interest corresponding to a polynucleotide capable of expressing a molecule, which interacts with a fused target ligand coupled with an enzyme or a fragment thereof.

According to another aspect, the invention also concerns a method of screening for a substance capable of stimulating
35 or inhibiting the interaction between a target ligand and a



molecule of interest, wherein the stimulating or the inhibiting activity is detected with a signal amplification system according to the invention, by means of generating a signal amplification and triggering transcriptional activation, and wherein said signal amplification and said triggering transcriptional activation are compared with those obtained from an identical signal amplification system without any substance.

The method of screening for substance capable of stimulating or inhibiting the interaction between a target ligand and a molecule of interest allows the choice of a substance acting positively or negatively or even not acting in this interaction.

In the method of screening for a substance capable of stimulating the interaction between a target ligand and a molecule of interest according to the invention, the signal amplification corresponds to the production of a signaling molecule and the transcriptional activation leads to a reporter gene expression.

In the method of screening for substance capable of inhibiting the interaction between a target ligand and a molecule of interest according to the invention, the signal amplification corresponding to the production of a signaling molecule is blocked or partially abolished and the transcriptional activation leading to a reporter gene expression is also blocked or partially abolished.

In one specific illustration of the present invention, the method of screening for a substance capable of stimulating or inhibiting the interaction between a target ligand and a molecule of interest, consists in a signal amplification system, which comprises a bacterial multi-hybrid system of at least two distinct fragments of an enzyme, whose enzymatic activity is restored by the interaction between the said molecule of interest and the said target ligand. The two fragments are any combination of



fragments from the enzyme, which *in vitro* functionally interact with the natural activator of said enzyme by restoring its activity.

According to the method of screening for substance
5 capable of stimulating or inhibiting the interaction between a target ligand and a molecule of interest of the present invention, the fragments are selected from the group consisting of :

10 (a) a fragment T25 corresponding to amino acids 1 to 224 of CyaA and a fragment T18 corresponding to amino acids 225 to 399 of CyaA;

(b) a fragment corresponding to amino acids 1 to 224 of CyaA and a fragment corresponding to amino acids 224 to 384 of CyaA;

15 (c) a fragment corresponding to amino acids 1 to 137 of CyaA and a fragment corresponding to amino acids 138 to 400 of CyaA;

(d) a fragment corresponding to amino acids 1 to 317 of CyaA and a fragment corresponding to amino acids 318 to 400
20 of CyaA; and

(e) two fragments from eukaryotic adenylate cyclase in association with molecules, such as G protein, and forskolin.

And more particularly, the fragments are a fragment T25
25 corresponding to amino acids 1 to 224 of *Bordetella pertussis* CyaA and a fragment T18 corresponding to amino acids 225 to 399 of *Bordetella pertussis* CyaA.

In another specific illustration of the present invention, the method of screening for a substance capable of
30 stimulating or inhibiting the interaction between a target ligand and a molecule of interest, consists in a signal amplification system, which comprises a bacterial multi-hybrid system of at least a first fragment of an enzyme and a modulating substance, whose activity, which is an enzymatic activity, is restored by the interaction between the said
35 molecule of interest and the said target ligand.



In both of the above illustrations of the present invention, the enzyme is selected from the group consisting of adenylate cyclase and guanylate cyclase from any origin, and more preferably the enzyme is the catalytic domain of
5 *Bordetella* adenylate cyclase (CyaA) located within the first 400 amino acid residues of the adenylate cyclase toxin.

The present invention further relates to a method of screening for a substance capable of stimulating or inhibiting the interaction between a target ligand and a
10 molecule of interest, wherein the substance is selected from the group consisting of protein, glycoprotein, lipoprotein, ligand, and any other drug having stimulating or inhibitory affinity.

The present invention also provides a kit for screening
15 for a substance capable of stimulating or inhibiting the interaction between a target ligand and a molecule of interest, wherein said kit comprises:

(a) a signal amplification system according to the invention with the substance capable of stimulating or
20 inhibiting the interaction between a target ligand and a molecule of interest;

(b) a signal amplification system according to the invention without any substance as the control;

(c) *E. coli* strain, or in any bacterial strain deficient
25 in endogenous adenylate cyclase, or any other eukaryotic cell; and

(d) a medium allowing for the detection of the complementation selected from the group consisting of indicator plate or selective medium as minimal medium
30 supplemented with lactose or maltose as unique carbon source, medium with antibiotics, medium to visualize fluorescence, conventional medium, and medium that allows the sorting by the presence of the phage receptor; and

(e) means for detecting whether the signal amplification
35 system with the substance is enhanced or inhibited with



respect to the signal amplification system without any substance.

The present invention includes a substance capable of stimulating or inhibiting the interaction between a target
5 ligand and a molecule of interest identified by the method of screening for a substance capable of stimulating or inhibiting the interaction between a target ligand and a molecule of interest according to the present invention.

According to a preferred embodiment of the invention,
10 the selection and the screening are performed in an *E. coli* strain, or in any bacterial strain deficient in endogenous adenylate cyclase, or any other eukaryotic cell.

Functional analysis of *B. pertussis* adenylate cyclase activity can be easily monitored in an *E. coli* strain
15 deficient in endogenous adenylate cyclase. In *E. coli*, cAMP bound to the transcriptional activator, CAP (catabolite activator protein), is a pleiotropic regulator of the expression of various genes, including genes involved in the catabolism of carbohydrates, such as lactose or maltose
20 (Ullmann, A. & Danchin, A. (1983) in *Advances in Cyclic Nucleotide Research* (Raven Press, New York), Vol. vol. 15, pp. 1-53). Hence, *E. coli* strains lacking cAMP are unable to ferment lactose or maltose. When the entire catalytic domain of CyaA (amino acids 1 to 399) is expressed in *E. coli cya*
25 under the transcriptional and translational control of *lacZ* (plasmid pDIA5240), its calmodulin-independent residual activity is sufficient to complement an adenylate cyclase deficient strain and to restore its ability to ferment lactose or maltose (Ladant, D., Glaser, P. & Ullmann, A.
30 (1992) *J. Biol. Chem.* 267, 2244-2250). This can be scored either on indicator plates (i.e. LB-X-Gal or MacConkey media supplemented with maltose) or on selective media (minimal media supplemented with lactose or maltose as unique carbon source).



The fact that the genetic tests according to the invention are carried out in *E. coli* greatly facilitates the screening as well as the characterization of the interaction between the target ligand and the molecule of interest.

5 Firstly, it is possible to use the same plasmid constructs to screen a library to identify the molecule of interest, also called a putative binding partner, to the target ligand, also called a given "bait", and then to express the target ligand and the molecule of interest in order to characterize their
10 interaction by *in vitro* binding assays.

Secondly, the high efficiency of transformation that can be achieved in *E. coli*, allows the analysis of libraries of high complexity. This is particularly useful for i) the screening and the selection of peptides from a library made
15 from random DNA sequences that present an affinity for a given bait protein, and ii) the exhaustive analysis of the network of interactions between the proteins of a given organism (Bartel, P. L., Roecklein, J. A., SenGupta, D. & Fields, S. (1996) *Nature Genetics* 12, 72-7; Fromont, R. M.,
20 Rain, J. C. & Legrain, P. (1997) *Nature Genetics* 16, 277-82).

The present invention further relates to a polynucleotide sequence coding for a signal amplification system according to the invention, wherein the polynucleotide sequence codes for a bacterial multi-hybrid system of at
25 least two chimeric polypeptides containing:

(a) a first chimeric polypeptide corresponding to a first fragment of an enzyme fused to a molecule of interest; and

(b) a second chimeric polypeptide corresponding to a
30 second fragment of an enzyme or a modulating substance capable of activating said enzyme fused to a target ligand.

The present invention also relates to a polynucleotide sequence coding for the signal amplification system according to the invention, wherein the polynucleotide sequence codes
35 for a bacterial multi-hybrid system containing:



(a) a first chimeric polypeptide corresponding to a first fragment of an enzyme fused to a molecule of interest;

(b) a second chimeric polypeptide corresponding to a second fragment of an enzyme or a modulating substance capable of activating said enzyme fused to a target ligand; and

(c) a substance capable of stimulating or inhibiting the interaction between a target ligand and a molecule of interest.

This invention will be described in greater detail with reference to the following examples.

Example 1

DHP1 is an adenylate cyclase deficient (*cya*) derivative of DH1 (F-, *glnV44(AS)*, *recA1*, *endA1*, *gyrA96 (Nal^r)*, *thi1*, *hsdR17*, *spoT1*, *rfbD1*) (25), and was isolated using phosphomycin as a selection antibiotic (Alper, M. D. & Ames, B. N. (1978) *J. Bacteriol.* 133, 149-57). Growth media used were the rich medium LB or the synthetic medium M63 (Miller, J. H. (1972) *Experiments in molecular genetics* (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.)) supplemented with 1 % carbon source. Antibiotic concentrations were as follows: ampicillin 100 mg/ml and chloramphenicol 30 mg/ml. Screening for the ability to ferment sugars was performed either on MacConkey agar plates containing 1 % maltose, or on LB plates containing 40 mg/ml X-Gal (5-Bromo-4-chloro-3-indolyl-b-D-galactopyranoside) and 0.5 mM IPTG (Isopropyl-b-D-thiogalactopyranoside).

Example 2

Plasmid pKT25 (3445-bp) is a derivative of the low copy vector pSU40 (expressing a kanamycin resistance selectable marker) that encodes the T25 fragment. It was constructed as follows: a 1044-bp *HindIII*-*EcoRI* fragment of pT25 was first subcloned into pSU40 linearized with *HindIII* and *EcoRI*, resulting in pKT25L. pKT25 was generated from pKT25L by deleting a 236-bp *NheI*-*HindIII* fragment.



Plasmid pUT18 (3023-bp) is a derivative of the high copy number vector pUC19 (expressing an ampicillin resistance selectable marker and compatible with pT25 or pKT25) that encodes the T18 fragment (amino acids 225 to 399 of CyaA). In a first step, we constructed plasmid pUC19L by inserting a 24-bp double-stranded oligonucleotide (5'-ATTCATCGATATACTAAGTAA-3' [SEQ ID No.: 1]) and its complementary sequence) between the *EcoRI* and *NdeI* sites of pUC19. Then, a 534-bp fragment harboring the T18 open reading frame was amplified by PCR (using appropriate primers and pT18 as target DNA) and cloned into pUC19L digested by *EcoRI* and *ClaI* (the appropriate restriction sites were included into the PCR primers). In the resulting plasmid, pUT18, the T18 open reading frame is fused in frame downstream of the multicloning site of pUC19. This plasmid is designed to create chimeric proteins in which a heterologous polypeptide is fused to the N-terminal end of T18 (see map).

Plasmid pUT18C (3017-bp) is a derivative of pUC19 (expressing a ampicillin resistance selectable marker and compatible with pT25 or pKT25) that encodes the T18 fragment. It was constructed by subcloning the same 534-bp PCR-amplified fragment harboring the T18 open reading frame described above into pUC19L linearized by *HindIII* and *PstI* (the appropriate restriction sites were included into the PCR primers). In the resulting plasmid, pUT18C, the T18 open reading frame is fused in frame upstream of the multicloning site of pUC19L. This plasmid is designed to create chimeric proteins in which a heterologous polypeptide is fused to the C-terminal end of T18 (see map).

Plasmid pKT25-zip (3556-bp) is a derivative of pKT25 that was constructed by inserting a DNA fragment (PCR-amplified using appropriate primers) encoding the leucine zipper region of GCN4 into pKT25 cleaved by *KpnI*, as described above.

Plasmid pUT18-zip (3125-bp) is a derivative of pUT18 that was constructed by inserting a 114bp DNA fragment (PCR-amplified using appropriate primers) encoding the leucine



zipper region of GCN4 into pUT18 linearized by KpnI and EcoRI.

Plasmid pUT18C-zip (3119-bp) is a derivative of pUT18C that was constructed by inserting the same 114-bp DNA fragment encoding the GCN4 leucine zipper described above into pUT18 linearized by KpnI and EcoRI.

Example 3

Briefly, a cAMP-biotinylated-BSA conjugate was coated on ELISA plates and non-specific protein binding sites were blocked with BSA. Boiled bacterial cultures were then added, followed by diluted rabbit anti-cAMP antiserum in 50 mM Hepes pH 7.5, 150 mM NaCl, 0.1 % Tween 20 (HBST buffer) containing 10 mg/ml BSA. After overnight incubation at 4°C, the plates were washed extensively with HBST, then goat anti-rabbit IgG coupled to alkaline phosphatase (AP) was added and incubated for 1 hr at 30°C. After washing, the AP activity was revealed by 5'-para-nitrophenyl phosphate. cAMP concentrations were calculated from a standard curve established with known concentrations of cAMP diluted in LB medium.

Example 4

Two compatible plasmids (derived from pACYC184 and pBluescript-II-KS) that express either the T25 fragment corresponding to amino acids 1 to 224 of CyaA or the T18 fragment corresponding to amino acids 225 to 399 were constructed. A multicloning site was fused to the C-terminal end of T25 to facilitate construction of fusions with foreign proteins. Similarly, the T18 fragment was fused in frame to *alacZ* of pBluescript-II-KS downstream of its multicloning siteK (Figure 2).

The two plasmids, pT25 and pT18, were co-transformed in DHP1, a *cya* derivative of the *E. coli* strain DH1 (Hanahan, D. (1983) *J. Mol. Biol.* 166, 557-80), and plated on MacConkey agar supplemented with maltose. As expected, no spontaneous complementation between the two isolated (independently expressed) fragments could be detected *in vivo*: all the



transformants were white (see Table 1). When the DHP1 strain was transformed with a plasmid expressing the full catalytic domain, all colonies were red (Table 1).

5 To test whether functional complementation between T25 and T18 could be brought about by fusing them to interacting proteins, there was inserted, within the multicloning site of both pT25 and pT18, a DNA sequence that codes for a 35 amino acid long leucine zipper derived from protein GCN4, a yeast transcriptional activator (Blondel, A. & Bedouelle, H. (1991) 10 *Protein Engineering* 4, 457-61). When the resulting plasmids, pT25-zip and pT18-zip, were co-transformed in DHP1 and plated on MacConkey/maltose media, the resulting colonies became red after 24-30 hours of growth at 30°C (Table 1).

15 Control experiments were carried out in which pT25-zip was co-transformed with pT18, or pT18-zip was co-transformed with pT25. None of the transformants exhibited complementation, demonstrating that the functional complementation of T25-zip and T18-zip was mediated by the interaction of their leucine zipper motif. The efficiency of 20 complementation could be further quantified by measuring in liquid cultures, either cAMP levels or β -galactosidase activities (Table 1).

Adenylate cyclase activities of the different transformants were measured in cell extracts in the presence 25 of CaM that binds tightly to T25 and T18 fragments to form the active adenylate cyclase complex. As shown in Table 1, only the extract from DHP1/pT25-zip/pT18-zip exhibited a significant enzymatic activity. The lack of activity in the extracts of the three other types of transformants indicates 30 that, at least one of the two complementary fragments of adenylate cyclase was missing, most probably as a consequence of its *in vivo* proteolytic degradation. Therefore, it would appear that the association of T25-zip and T18-zip, through their leucine zipper motif, not only resulted in their



functional complementation, but also in their stabilization. Stabilization of protein fragments (a and w peptides) through complementation (Ullmann, A., Jacob, F. & Monod, J. (1968) *J. Mol. Biol.* 32, 1-13) has also been observed for β -
S galactosidase.



Table 1: Analysis of complementation in DHP1 strain

Plasmids	Phenotype on Mac Conkey/maltose	β -galactosidase ^(a)	cAMP ^(b)	Adenylate cyclase activity ^(c)	
				+ CaM	-CaM
none	White	179	< 10	< 1	< 0.01
pCm-AHL1	Red / 24 hrs	6650	3400	13,000	10
pT25 + pT18	White / 72 hrs	130	< 10	< 1	< 0.01
pT25 + pT18-zip	White / 72 hrs	183	< 10	< 1	< 0.01
pT25-zip + pT18	White / 72 hrs	178	< 10	< 1	< 0.01
pT25-zip + pT18-zip	Red / 30 hrs	4750	1100	10,000	4

34

a) units/mg dry weight bacteria

b) pmol/mg dry weight bacteria

c) nmol cAMP protein; when present in the assays, CaM was at a concentration of 1 micromolar. Bacteria were grown in LB at 30°C in the presence of 0.5 mM IPTG plus appropriate antibiotics. The results represent the average values obtained for at least five independent cultures, which differed by less than 10 %.



Example 5

Screening for *in vivo* protein-protein interactions by using functional complementation of T25 and T18 was carried out.

5 The goal was to examine whether the complementation between T25 and T18 could be used to analyze interactions between proteins larger than the 35-residue long leucine zipper motif. A DNA fragment that encodes the N-terminal part (residues 1 to 290) of the dimeric tyrosyl tRNA synthetase
10 from *Bacillus stearothermophilus* (Guez-Ivanier, V. & Bedouelle, H. (1996) *J. Mol. Biol.* 255, 110-120) was subcloned into the multicloning site of plasmids pT25 and pT18. The resulting plasmids, pT25-TyrRS and pT18-TyrRS, when co-transformed in DHP1, yielded red transformants on
15 MacConkey/maltose. The transformants synthesized cAMP and expressed β -galactosidase (Table 2). Control transformations confirmed that the TyrRS moiety was responsible for the functional complementation between T25-TyrRS and T18-TyrRS. Furthermore, no complementation occurred when T25-TyrRS was
20 cotransformed with pT18-zip or vice versa. This demonstrates that the complementation was dictated by the specificity of recognition of the polypeptides fused to the two fragments, T25 and T18.

 It was further shown (Table 2) that the bacterial multi-
25 hybrid system could detect interaction between the yeast splicing factors Prp11 and Prp21 (fused to T25 and T18, respectively) that was previously characterized in the yeast two-hybrid assay (Legrain, P. & Chapon, C. (1993) *Science* 262, 108-10). This demonstrates that this bacterial
30 complementation assay can reveal association between eukaryotic proteins.



Table 2: Complementation between various chimeric proteins

Plasmids	Phenotype on Mac Conkey/maltose	β -galactosidase ^(a)	CAMP ^(b)	
pT25-Tyr + pT18-Tyr	Red / 40 hrs	2800	580	580
pT25-Tyr + pT18	White / 96 hrs	193	< 10	10
pT25 + pT18-Tyr	White / 96 hrs	183	< 10	< 0.01
pT25 - Tyr + pT18-zip	White / 96 hrs	134	< 10	< 0.01
pT25-prp11+ pT18-prp21	Red / 40 hrs	850	65	4
				36

a) units/mg dry weight bacteria

b) pmol/mg dry weight bacteria

Bacteria were grown in LB at 30°C in the presence of 0.5 mM IPTG plus appropriate antibiotics. The results represent the average values obtained for at least five independent cultures.



To mimic a screening procedure, plasmids pT18-zip and pT18-TyrRS were mixed with about a 5-fold excess of pT18 and co-transformed this mixture in DHP1 with either pT25 or pT25-zip. The transformants were plated on LB-X-Gal. All the colonies co-transformed with pT25 were white (Fig. 4A). Around 20 % of the colonies were blue when the cells were co-transformed with the mixture of pT18 derivatives and pT25-zip. The plasmid DNAs of these clones were further analyzed by restriction mapping. As expected, the blue colonies among the bacteria co-transformed with pT25-zip harbored only pT18-zip (Fig. 4B).

In another series of experiments, pT18-zip was mixed with a 1,000 fold excess of pT18 and this mixture was transformed in DHP1 harboring pT25-zip and plated on MacConkey/maltose. Three red colonies were identified among about 3,000 white ones. Plasmid DNA analysis of the Mal⁺ clones confirmed the presence of pT18-zip. Transformation of the same mixture of pT18-zip/pT18 into DHP1 harboring pT25 gave no Mal⁺ clones out of 10,000 analyzed (data not shown). These results indicate that the functional complementation between the adenylate cyclase fragments could be used to identify interacting proteins in *E. coli*.

Finally, an examination was made to determine whether the complementation between T25 and T18 could be used in a selection procedure rather than using the screening described above. DHP1 bacteria cotransformed with complementing plasmids (pT25-zip/pT18-zip or pT25-TyrRS/pT18-TyrRS) were able to grow on minimal media supplemented with lactose or maltose as unique carbon sources, while bacteria cotransformed with non-complementing plasmids (pT25-zip/pT18-TyrRS or pT25-TyrRS/pT18-zip) did not grow.

To determine whether this selection could be used to identify interacting proteins among an excess of non-interacting ones, the following "model screening" was performed on selective media: DHP1 bacteria harboring pT25-



zip and pT18-zip (expected phenotype: Lac⁺) were mixed with a 10⁵-excess of DHP1/pT25/pT18 (expected phenotype: Lac⁻), and then 10⁷ cells from this mixture were plated on minimal media supplemented with lactose plus antibiotics. After 4-5 days at 30°C, 100 to 200 Lac⁺ colonies appeared. Plasmid DNA analysis indicated that 18 out of 20 of these colonies tested harbored pT25-zip and pT18-zip. When 10⁷ DHP1/pT25/pT18 cells were plated on minimal media/lactose, about 10 colonies were detected: these cells appeared to represent spontaneous revertants of DHP1 to a Lac⁺ phenotype (due to either reversion of *cya*⁻ to *cya*⁺ or to cAMP/CAP independent *lac* promoter mutations). This "model screening" demonstrates that bacteria expressing specific interacting proteins fused to the adenylate cyclase fragments could be selected among a large number (here a 10⁵-fold excess) of irrelevant clones.

Example 6

A further test was carried out to determine whether functional complementation could be obtained when the interacting polypeptide is fused at the C-terminus of T18 rather than at its N-terminus. Two new plasmids derived from the pUC19 vector were constructed for this purpose. In pUT18, the T18 polypeptide is fused in frame downstream to the multicloning site of pUC19, whereas in pUT18C, the T18 polypeptide is fused in frame upstream of the multicloning site. A DNA fragment encoding the leucine zipper of GCN4 was then cloned in frame into both pUT18 and pUT18C to yield pUT18-zip and pUT18C-zip.

As shown in Table 3, co-transformation of DHP1 with pT25-zip and either pUT18-zip or pUT18C-zip led to functional complementation. This indicates that interacting polypeptides could be fused at both ends of the T18 fragment with the same complementation efficiency. Similar results were obtained (Table 3) when the T25-zip chimeric protein was expressed from a pSU40 derivative that expresses the kanamycin resistance selectable marker (pKT25-zip).



Table 3: Comparison of complementation between N-terminal and C-terminal fusion proteins

Plasmids	Phenotype on MacConkey/maltose	β -galactosidase units/mg dry weight bacteria
pT25 + pT18	White / 72 hrs	154
pT25-zip + pUT18-zip	Red / 26 hrs	5100
pT25-zip + pUT18	White / 72 hrs	ND
pT25 + pUT18-zip	White / 72 hrs	ND
pT25-zip + pUT18-zip	Red / 26 hrs	6180
pT25-zip + pUT18C	White / 72 hrs	ND
pT25 + pUT18C-zip	White / 72 hrs	ND
pT25-zip + pUT18C-zip	Red / 26 hrs	6100
pKT25-zip + pUT18	White / 72 hrs	ND
pKT25 + pUT18-zip	White / 72 hrs	ND
pKT25-zip + pUT18-zip	Red / 26 hrs	ND
pKT25-zip + pUT18C	White / 72 hrs	ND
pKT25 + pUT18C-zip	White / 72 hrs	ND
pKT25-zip + pUT18C-zip	Red / 26 hrs	ND

5 Bacteria were grown in LB at 30°C in the presence of appropriate antibiotics.

ND: not done.

Example 7

10 The bacterial two-hybrid system was used to analyze interactions between various sub-domains of the dimeric tyrosyl-tRNA synthetase from *B. stearothermophilus*, (TyrRS),



which is a symmetrical dimer (Brick, P. and D. M. Blow (1987) *J Mol. Biol.* 194,2 87-297). Its monomer is composed of two domains.

5 Deletion of the C-terminal domain (321-419) of TyrRS produces a truncated Δ TyrRS which activates tyrosine as the full-length molecule but is no longer able to bind the cognate tRNA. The truncated Δ TyrRS (1-320) forms a dimer that closely resembles the wild-type one.

10 The crystal structure of the dimeric Δ TyrRS revealed that each monomer contains two structural domains: an α/β domain (1-220) containing six-stranded β -sheets and an α -helical domain (221-320) containing five helices (Brick, P. and D. M. Blow 1987) *J. Mol. Biol.* 194, 287-297). The dimer is formed by the association of a hydrophobic surface
15 encompassing residues 128-167 within the α/β domain of each subunit.

To analyze interactions between various sub-domains, different fragments of the TyrRS polypeptide (generated by PCR using appropriate primers) were fused in frame with
20 either the T25 or the T18 fragment and the resulting chimeric proteins were tested for functional complementation in DPH1. The results, summarized in Fig. 5, revealed 3 different types of interactions between the TyrRS monomers or between the TyrRS sub-domains:

25 1) Dimerisation through the α/β domains as can be seen in the crystal structure of TyrRS. For instance, the chimeric protein T18-TyrRS1-249, which harbors only the α/β domain, can fully complement T25-TyrRS1-333, which contains both the α/β and α domains. Previous studies have shown that
30 introduction of charged residues into the hydrophobic subunit interface of dimeric TyrRS induces reversible dissociation (Ward, W.H.J., H. Jones, and A.R. Ferscht (1987) *Biochemistry* 26, 4131-4138). Confirmation that a point mutation that



converts Phe164 to Arg abolishes the interaction between the α/β domains (as shown by the absence of functional complementation) has been made.

2) Dimerisation through the α domains, which has not
5 been previously predicted from the crystal structure of TyrRS. Analysis of complementation between various fragments indicated that this dimerisation is mediated by the C-terminal region of the (α domain (Fig. 5)). This region contains a pseudo leucine zipper motif (LLL on Fig. 5) made
10 of 3 leucine residues at positions 298, 305, and 312. The same segment can also mediate a specific interaction with the GCN4 leucine-zipper (Fig. 5).

3) Interaction between the α/β domain and the α domain.

This study illustrates the interest of the bacterial
15 two-hybrid in delineating interacting domains of proteins and shows that it could reveal interactions, which occur *in vivo* and that were not expected from the three dimensional structure.

Example 8

20 The two-hybrid system was also used to analyze the dimerization of a DNA-binding protein from *B. pertussis*, BvgA. BvgA is a transcriptional regulator, which in *B. pertussis* controls the expression of virulence-associated genes. It is a member of the bacterial two-component signal transduction
25 family, together with its cognate sensor protein, BvgS (Scarlato et al. (1990) *Proc. Natl. Acad. Sci. USA*, 87:6753). The transmembrane BvgS is autophosphorylated in response to environmental signals and subsequently phosphorylates BvgA. BvgA, in its phosphorylated form, can bind to specific DNA
30 sequences within the promoter of several virulence genes and activates their transcription. Several studies previously suggested that BvgA might have the capacity to dimerize although a direct demonstration that BvgA is a dimer is still lacking.



To study the dimerization of BvgA, a set of plasmids was constructed that encode various fragments of the BvgA polypeptide fused to either T25 or T18. These plasmids were co-transformed in DHP1 (Fig. 6), and the level of functional complementation between the different chimeric proteins was determined by measuring β -galactosidase activities. These results indicate that BvgA can indeed dimerize and that the critical region required for dimerization is localized within the central part of the protein.

10 Example 9

A selection procedure was also established that will permit an easy screening for mutations that abolish the interaction between two hybrid proteins. This selection is based on the well established fact that *E.coli cya*⁺ strains are resistant to phage λ , whereas *E.coli cya* strains are sensitive. The phage receptor, the LamB protein, is the product of the *lamB* gene, which is part of the maltose regulon; therefore, its expression requires cAMP. In consequence, cells producing cAMP will lyse when infected with λ vir. Molecules or mutations that abolish the interaction between two hybrid proteins will abrogate cAMP synthesis, and, therefore, the cells should become resistant to phage λ .

Quantification of resistance to phage λ vir has been performed in liquid Luria broth. Experimental conditions of infection that enable complete lysis of cAMP producing bacteria are the following:

- multiplicity of infection: 2 to 10,
- MgSO₄ concentration: 20 mM,
- incubation time at 37°C under aeration for 2 to 3 hours.

DHP1 bacteria were grown overnight in the presence of 1 mM cAMP in Luria broth. The bacteria were washed 3 times with Luria broth, and an aliquot was immediately infected with



λ vir. Another aliquot was diluted and cultivated in the absence of cAMP at 37°C. After 15 generations of growth, a sample of the bacteria grown in the absence of cAMP has been infected with λ vir. After serial dilutions, bacteria were plated on solid Luria broth and counted. Out of 1.6×10^9 bacteria grown in the absence of cAMP 1.2×10^9 phage resistant clones were counted (75 %), whereas out of 1.3×10^9 bacteria grown in the presence of cAMP, only 140 λ -resistant clones (frequency of 10^{-9}) were found. These latter λ -resistant clones were white on MacConkey maltose plates, suggesting that they were cAMP-independent *malT* mutants.

Plasmids useful for practicing this invention have been deposited at Collection Nationale de Cultures de Microorganismes in Paris, France on November 25, 1998, as follows:

<u>Plasmid</u>	<u>Accession No.</u>
XL-1/pUT18	I-2092
XL-1/pUT18C	I-2093
XL-1/pT25	I-2094
XL-1/pKT25	I-2095

As it appears from the teachings of the specification, the invention is not limited in scope to one or several of the above detailed embodiments; the present invention also embraces all the alternatives that can be performed by one skilled in the same technical field, without deviating from the subject or from the scope of the instant invention.



CLAIMS

1. A signal amplification system comprising a bacterial multi-hybrid system of at least two chimeric polypeptides
5 containing:

(a) a first chimeric polypeptide corresponding to a first fragment of an enzyme;

(b) a second chimeric polypeptide corresponding to a second fragment of an enzyme or a modulating substance
10 capable of activating said enzyme,

wherein the first fragment is fused to a molecule of interest and the second fragment or the modulating substance is fused to a target ligand and wherein the activity of the enzyme is restored by the *in vivo* interaction between the said molecule
15 of interest and the said target ligand and wherein a signal amplification is generated.

2. The signal amplification system according to claim 1, wherein the enzyme is an enzyme selected from the group consisting of adenylate cyclase and guanylate cyclase from
20 any origin.

3. The signal amplification system according to claim 2, wherein the enzyme is the catalytic domain of *Bordetella* adenylate cyclase (CyaA), located within the first 400 amino acid residues of the adenylate cyclase toxin.

4. The signal amplification system according to claim 3, wherein the first and the second fragments are any combination of fragments from the same enzyme which *in vitro* functionally interact with the natural activator of said enzyme by restoring its activity.

5. The signal amplification system according to claim 4, wherein the first and the second fragments are selected from the group consisting of :

(a) a fragment T25 corresponding to amino acids 1 to 224 of CyaA and a fragment T18 corresponding to amino acids 225
35 to 399 of CyaA;



(b) a fragment corresponding to amino acids 1 to 224 of CyaA and a fragment corresponding to amino acids 224 to 384 of CyaA;

5 (c) a fragment corresponding to amino acids 1 to 137 of CyaA and a fragment corresponding to amino acids 138 to 400 of CyaA;

(d) a fragment corresponding to amino acids 1 to 317 of CyaA and a fragment corresponding to amino acids 318 to 400 of CyaA;

10 (e) two fragments from eukaryotic adenylate cyclase in association with molecules such as G protein, forskolin.

6. The signal amplification system according to claim 4 or 5, wherein the first and the second fragments are a fragment T25 corresponding to amino acids 1 to 224 of
15 *Bordetella pertussis* CyaA and a fragment T18 corresponding to amino acids 225 to 399 of *Bordetella pertussis* CyaA.

7. The signal amplification system according to any one of the claims 1 to 3, wherein the modulating substance is a natural activator, or a fragment thereof, of the enzyme.

20 8. The signal amplification system according to claim 7, wherein the natural activator is the calmodulin (CaM), or a fragment thereof, and said first fragment is mutated compared to the wild type enzyme.

25 9. The signal amplification system according to claim 8, wherein the first fragment is a mutated fragment of the catalytic domain of *Bordetella* adenylate cyclase (CyaA).

30 10. A method of selecting a molecule of interest which is capable of binding to target ligand wherein the interaction between the said molecule of interest and the said target ligand is detected with a signal amplification system according to any one of the claims 1 to 9, by means of generating a signal amplification and triggering transcriptional activation.



11. The method of selecting a molecule of interest according to claim 10, wherein the signal amplification corresponds to the production of a signaling molecule.

12. The method of selecting a molecule of interest
5 according to claim 10, wherein the transcriptional activation leads to a reporter gene expression.

13. The method of selecting a molecule of interest according to any one of claims 10 to 12, wherein the signal amplification system comprises a bacterial multi-hybrid
10 system of at least two distinct fragments of an enzyme, whose enzymatic activity is restored by the interaction between the said molecule of interest and the said target ligand.

14. The method of selecting a molecule of interest according to any one of claims 10 to 12, wherein the signal
15 amplification system comprises bacterial multi-hybrid system of at least a first fragment of an enzyme and a modulating substance, whose activity is restored by the interaction between the said molecule of interest and the said target ligand.

15. The method of selecting a molecule of interest according to any one of claims 10 to 14, wherein the target
20 ligand is selected from the group consisting of protein, peptide, polypeptide, receptor, ligand, antigen, antibody, DNA binding protein, glycoprotein, lipoprotein and recombinant protein.
25

16. The method of selecting a molecule of interest according to any one of claims 10 to 15, wherein the molecule of interest is capable of interacting with the target ligand and possibly of binding to said target ligand.

17. The method of selecting a molecule of interest according to any one of claims 10 to 16, wherein the
30 interaction between the molecule of interest and the target ligand is detected, by means of signal amplification which triggers transcriptional activation, and is quantified by



measuring the synthesis of the signaling molecule or the expression of the reporter gene.

18. The method of selecting a molecule of interest according to claim 11, wherein the signaling molecule
5 corresponds to the synthesis of cAMP.

19. The method of selecting a molecule of interest according to claim 11, wherein the signaling molecule corresponds to the synthesis of cGMP.

20. The method of selecting a molecule of interest
10 according to claim 12, wherein the reporter gene expression is selected from the group consisting of gene coding for nutritional marker such as lactose, maltose; gene conferring resistance to antibiotics such as ampicillin, kanamycin or tetracyclin; gene encoding for toxin; color marker such as
15 fluorescent marker of the type of the Green Fluorescent Protein (GFP); gene encoding for phage receptor proteins or fragment thereof such as phage λ receptor, *lamB* and any other gene giving a selectable phenotype.

21. The method of selecting a molecule of interest
20 according to any one of claims 10 to 20, wherein the molecule of interest is a mutant molecule compared to the known wild type molecule and said molecule of interest is tested for its capacity of interacting with the target ligand.

22. The method of selecting a molecule of interest
25 according to any one of claims 10 to 21, wherein the selection is performed in an *E. coli* strain, or in any bacterial strain deficient in endogenous adenylate cyclase or any other eukaryotic cell.

23. A kit for selecting molecule of interest, wherein
30 said kit comprises:

(a) a signal amplification system according to any one of claims 1 to 9;

(b) an *E. coli* strain, or in any bacterial strain deficient in endogenous adenylate cyclase or any other
35 eukaryotic cell and;



(c) a medium allowing the detection of the complementation selected from the group consisting of indicator or selective medium as minimal medium supplemented with lactose or maltose as unique carbon source, medium with antibiotics, medium to visualize fluorescence, conventional medium and medium which allows the sorting by the presence of the phage receptor.

24. A kit for selecting molecule of interest, wherein said kit comprises:

(a) a signal amplification system according to any one of claims 1 to 9 wherein the molecule of interest is a mutant molecule compared to the known wild type molecule;

(b) a signal amplification system according to any one of claims 1 to 9 wherein the molecule of interest is the known wild type molecule as the control;

(c) *E. coli* strain, or in any bacterial strain deficient in endogenous adenylate cyclase or any other eukaryotic cell and;

(d) a medium allowing the detection of the complementation selected from the group consisting of indicator or selective medium as minimal medium supplemented with lactose or maltose as unique carbon source, medium with antibiotics, medium to visualize fluorescence, conventional medium and medium which allows the sorting by the presence of the phage receptor for each signal amplification system;

(e) means for detecting whether the signal amplification system with the mutant molecule is enhanced or inhibited with respect to the signal amplification system with wild type.

25. A method of screening for substance capable of stimulating or inhibiting the interaction between a target ligand and a molecule of interest wherein respectively the stimulating or the inhibiting activity is detected with a signal amplification system according to any one of the claims 1 to 9, by means of generating an amplification and respectively of triggering or of abolishing transcriptional



activation, and wherein said signal amplification and said triggered or abolished transcriptional activation are compared with those obtained from an identical signal amplification system without any substance.

5 26. The method of screening for substance capable of stimulating or inhibiting the interaction between a target ligand and a molecule of interest according to claim 25, wherein the signal amplification system comprises a bacterial multi-hybrid system of at least two distinct fragments of an
10 enzyme, whose enzymatic activity is restored by the interaction between the said molecule of interest and the said target ligand.

 27. The method of screening for substance capable of stimulating or inhibiting the interaction between a target
15 ligand and a molecule of interest according to claim 25, wherein the signal amplification system comprises a bacterial multi-hybrid system of at least a first fragment of an enzyme and a modulating substance, whose activity is restored by the interaction between the said molecule of interest and the
20 said target ligand.

 28. The method of screening for substance capable of stimulating the interaction between a target ligand and a molecule of interest according to any one of claims 25 to 27, wherein the signal amplification corresponds to the
25 production of a signaling molecule.

 29. The method of screening for substance capable of inhibiting the interaction between a target ligand and a molecule of interest according to any one of claims 25 to 27, wherein the signal amplification corresponding to the
30 production of a signaling molecule is blocked or partially abolished.

 30. The method of screening for substance capable of stimulating the interaction between a target ligand and a molecule of interest according to any one of claims 25 to 28,



wherein the transcriptional activation leads to a reporter gene expression.

31. The method of screening for substance capable of inhibiting the interaction between a target ligand and a molecule of interest according to any one of claims 25 to 27 and to claim 29, wherein the transcriptional activation leading to a reporter gene expression is blocked or partially abolished.

32. The method of screening for substance capable of stimulating or inhibiting the interaction between a target ligand and a molecule of interest according to any one of claims 25 to 31, wherein the target ligand is selected from the group consisting of receptor, ligand, antigen, antibody, DNA binding protein, glycoprotein and lipoprotein.

33. The method of screening for substance capable of stimulating or inhibiting the interaction between a target ligand and a molecule of interest according to any one of claims 25 to 32, wherein the substance is selected from the group consisting of protein, glycoprotein, lipoprotein, ligand and any other drug having stimulating or inhibitory affinity.

34. The method of screening for substance capable of stimulating or inhibiting the interaction between a target ligand and a molecule of interest according to claim 28 or 29, wherein the signaling molecule corresponds to the synthesis of cAMP.

35. The method of screening for substance capable of stimulating or inhibiting the interaction between a target ligand and a molecule of interest according to claim 28 or 29, wherein the signaling molecule corresponds to the synthesis of cGMP.

36. The method of screening for substance capable of stimulating or inhibiting the interaction between a target ligand and a molecule of interest according to claim 30 or 31, wherein the reporter gene expression is selected from the



group consisting of gene coding for nutritional marker such as lactose, maltose; gene conferring resistance to antibiotics such as ampicillin, kanamycin or tetracyclin; gene encoding for toxin; color marker such as fluorescent marker of the type of the Green Fluorescent Protein (GFP);
5 gene encoding for phage receptor proteins or fragment thereof such as phage λ receptor, *lamB* and any other gene giving a selectable phenotype.

37. The method of screening for substance capable of stimulating or inhibiting the interaction between a target
10 ligand and a molecule of interest according to any one of claims 25 to 36, wherein the molecule of interest is a mutant molecule compared to the known wild type molecule and said molecule of interest is tested for its capacity of
15 interacting with the target ligand.

38. The method of screening for substance capable of stimulating or inhibiting the interaction between a target ligand and a molecule of interest according to any one of claims 25 to 37, wherein the screening is performed in an *E.*
20 *coli* strain, or in any bacterial strain deficient in endogenous adenylate cyclase or any other eukaryotic cell.

39. A kit for screening for substance capable of stimulating or inhibiting the interaction between a target ligand and a molecule of interest, wherein said kit
25 comprises:

(a) a signal amplification system according to any one of claims 1 to 9 with the substance capable of stimulating or inhibiting the interaction between a target ligand and a molecule of interest;

30 (b) a signal amplification system according to any one of claims 1 to 9 without any substance as the control;

(c) *E. coli* strain, or in any bacterial strain deficient in endogenous adenylate cyclase or any other eukaryotic cell and;



(d) a medium allowing the detection of the complementation selected from the group consisting of indicator plate or selective medium as minimal medium supplemented with lactose or maltose as unique carbon source, medium with antibiotics, medium to visualize fluorescence, conventional medium and medium which allows the sorting by the presence of the phage receptor and;

(e) means for detecting whether the signal amplification system with the substance is enhanced or inhibited with respect to the signal amplification system without any substance.

40. A molecule of interest identified by the method of any one of the claims 10 to 22.

41. A molecule of interest corresponding to a polynucleotide capable of expressing a molecule which interacts with a fused target ligand coupled with an enzyme or a fragment thereof.

42. A substance capable of stimulating or inhibiting the interaction between a target ligand and a molecule of interest identified by the method of any one of the claims 25 to 38.

43. The signal amplification system according to any one of the claims 1 to 9, wherein the bacterial multi-hybrid system contains:

(a) a first chimeric polypeptide corresponding to a first fragment a of an enzyme;

(b) a second chimeric polypeptide corresponding to a second fragment of an enzyme or a modulating substance capable of activating said enzyme and;

(c) a substance capable of stimulating or inhibiting the interaction between a target ligand and a molecule of interest, wherein the first fragment is fused to a molecule of interest and the second fragment or the modulating substance is fused to a target ligand and wherein the activity of the enzyme is restored by the interaction between



the said molecule of interest and the said target ligand and wherein a signal amplification is generated.

44. Polynucleotide sequence coding for the signal amplification system according to any one of the claims 1 to 9, wherein the polynucleotide sequence codes for a bacterial multi-hybrid system of at least two chimeric polypeptides containing:

(a) a first chimeric polypeptide corresponding to a first fragment a of an enzyme fused to a molecule of interest;

(b) a second chimeric polypeptide corresponding to a second fragment of an enzyme or a modulating substance capable of activating said enzyme fused to a target ligand.

45. Polynucleotide sequence coding for the signal amplification system according to any one of the claims 1 to 9 and to claim 43, wherein the polynucleotide sequence codes for a bacterial multi-hybrid system containing:

(a) a first chimeric polypeptide corresponding to a first fragment a of an enzyme fused to a molecule of interest;

(b) a second chimeric polypeptide corresponding to a second fragment of an enzyme or a modulating substance capable of activating said enzyme fused to a target ligand;

(c) a substance capable of stimulating or inhibiting the interaction between a target ligand and a molecule of interest.



FIGURE 1A

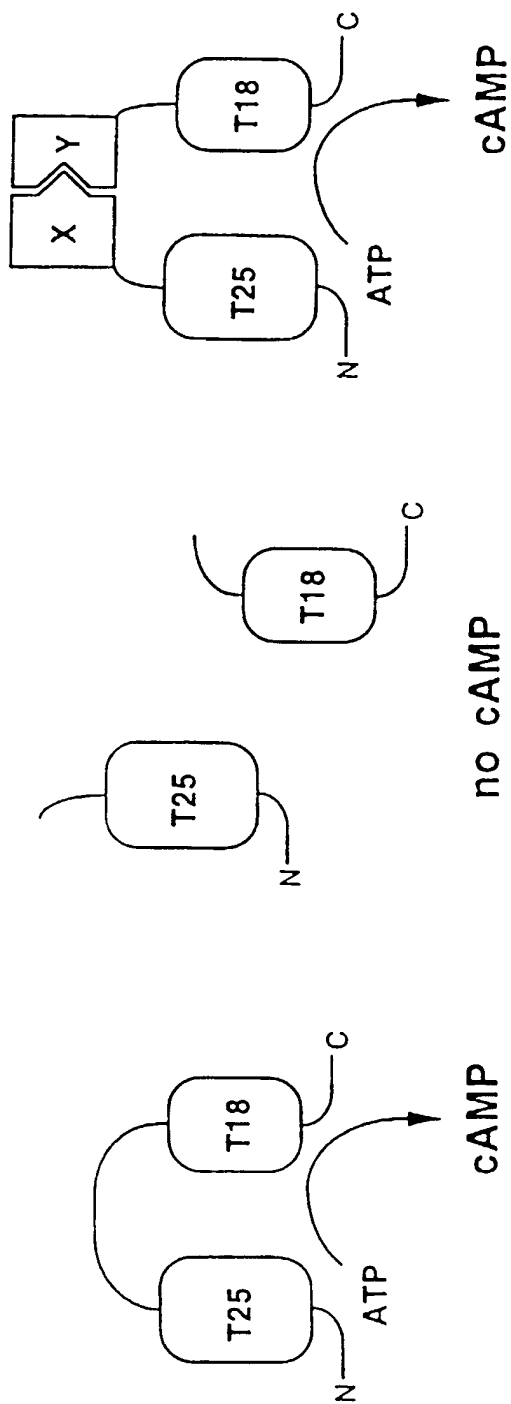


FIGURE 1B

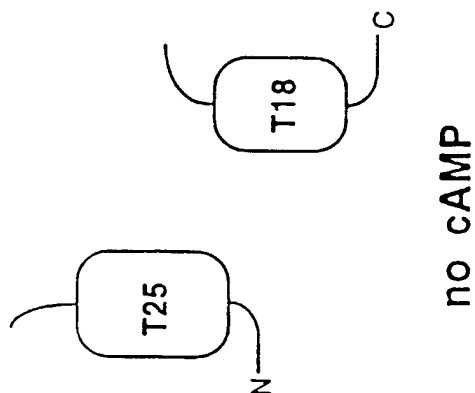


FIGURE 1C

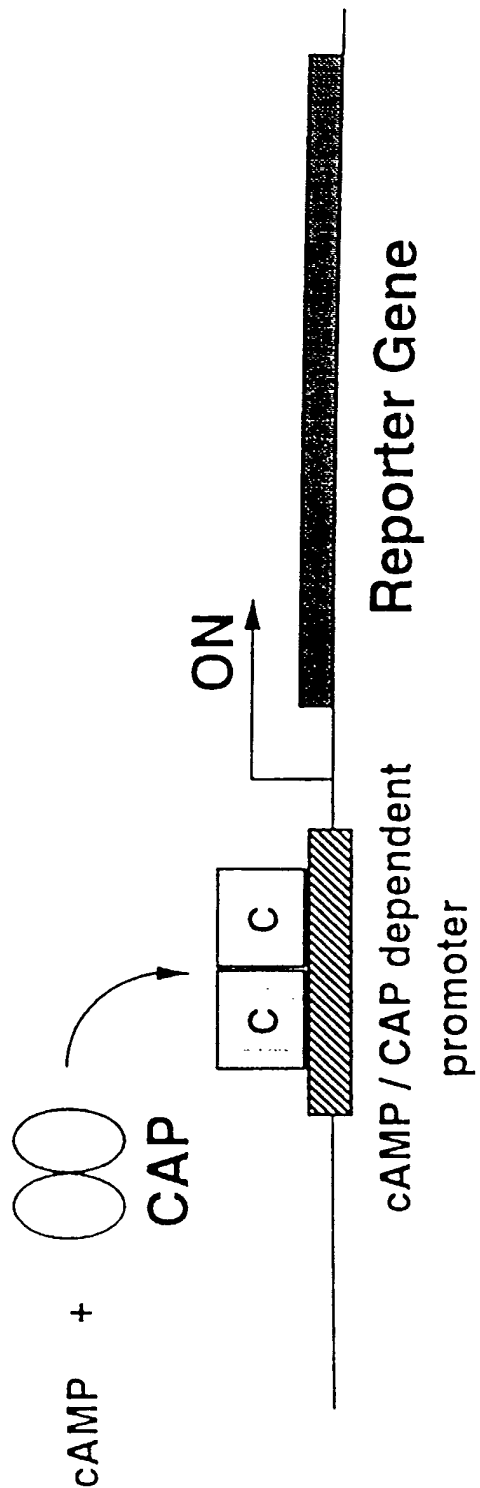
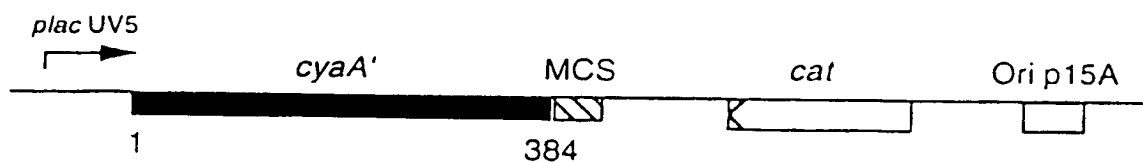


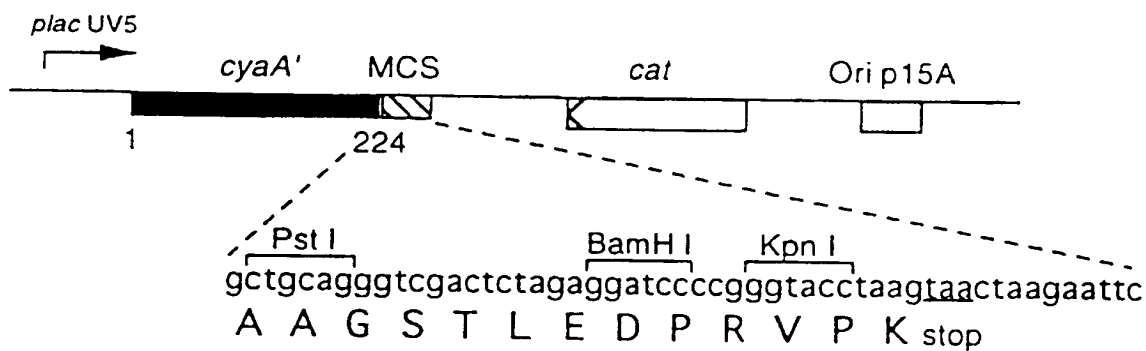
FIGURE 1D

10 Rec'd PCH-10 0 2 JUN 2000

pCmAHL1



pT25



pT18

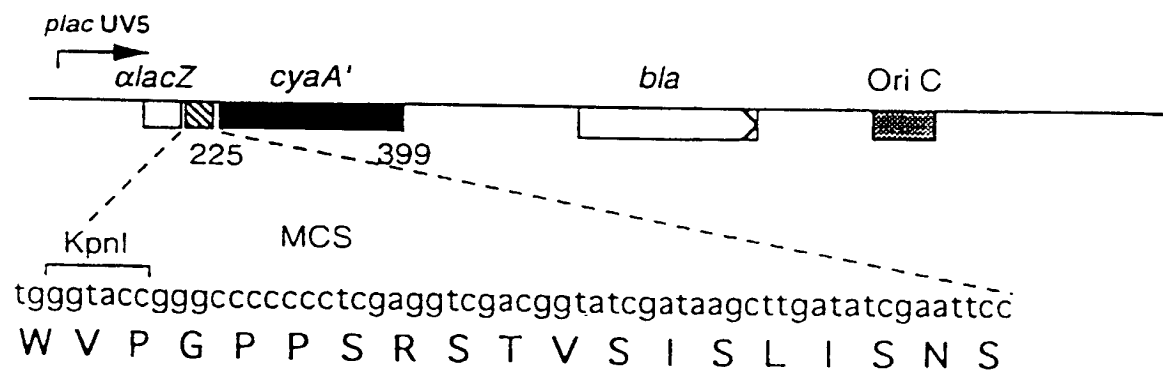
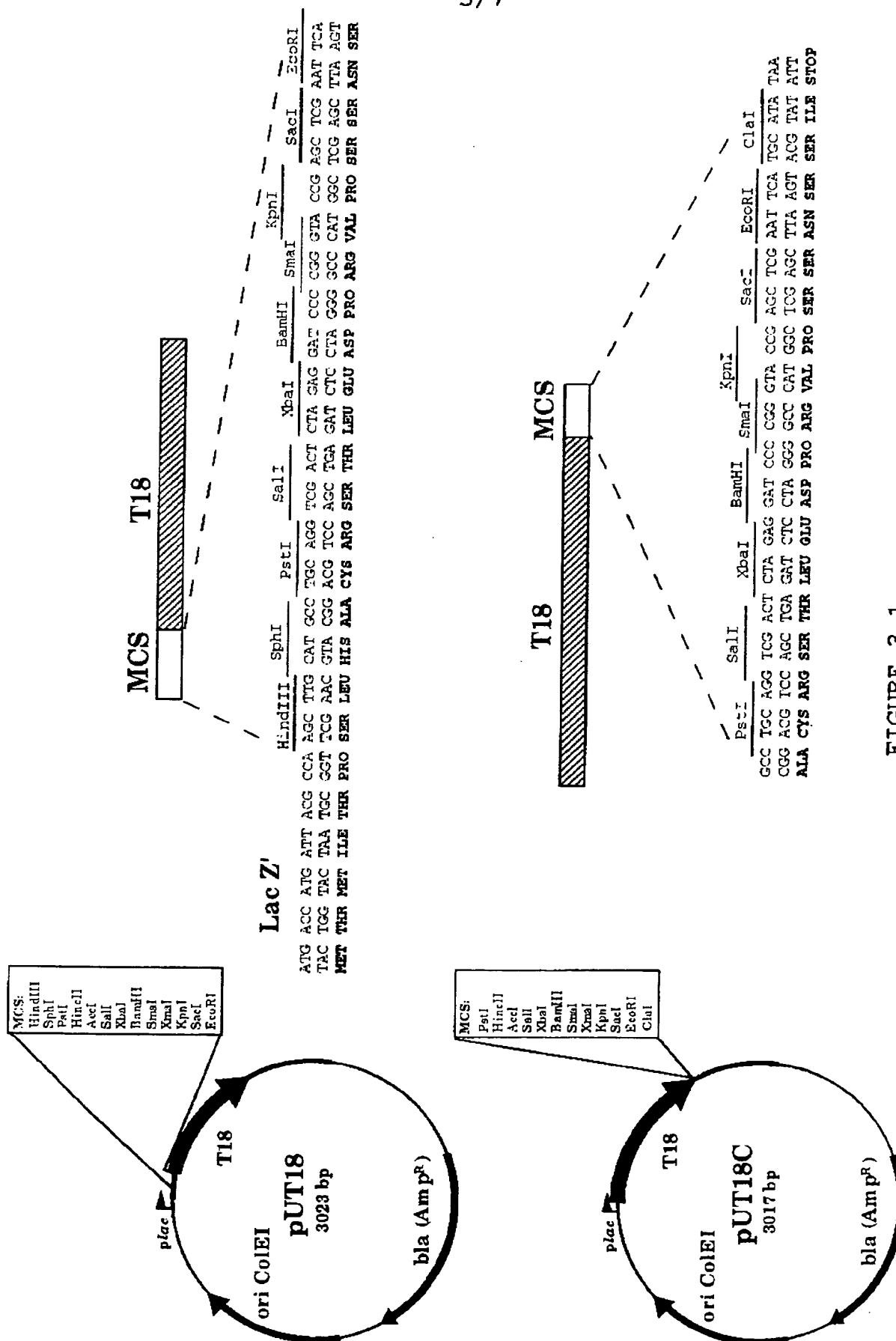


FIGURE 2

10 Rec'd PCWFS 0 2 JUN 2000

3/7

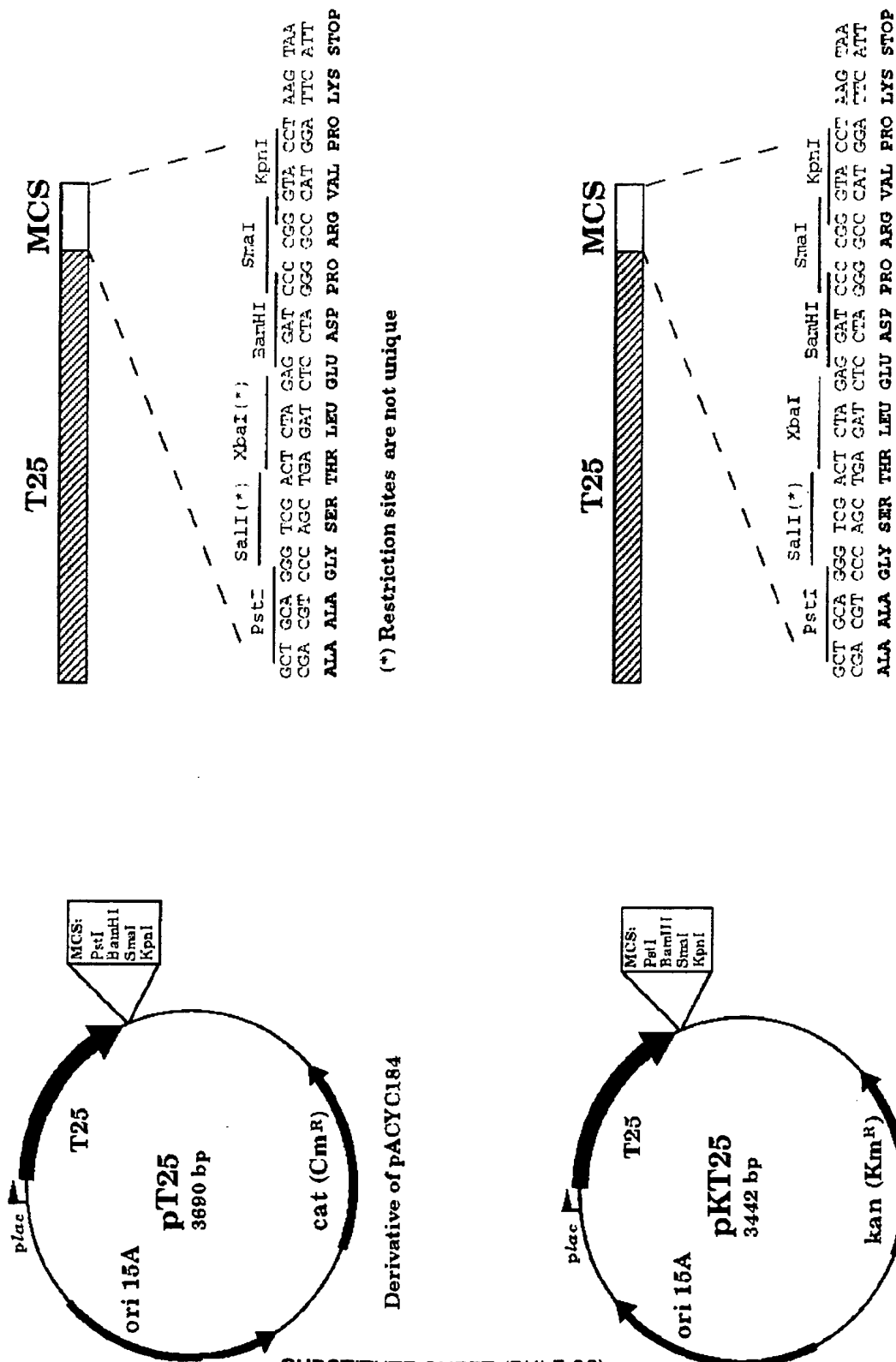
VECTORS EXPRESSING THE T18 FRAGMENT



10 Rec'd FCTH 10 02 JUN 2000

4/7

VECTORS EXPRESSING THE T25 FRAGMENT



(*) Restriction site is not unique

FIGURE 3.2

10 Rec'd POTTS 0 2 JUN 2000

5/7

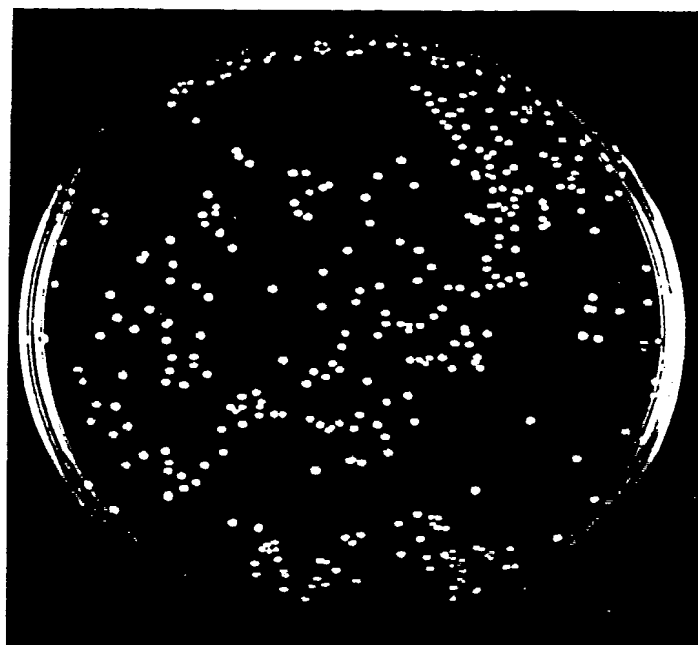


FIGURE 4A

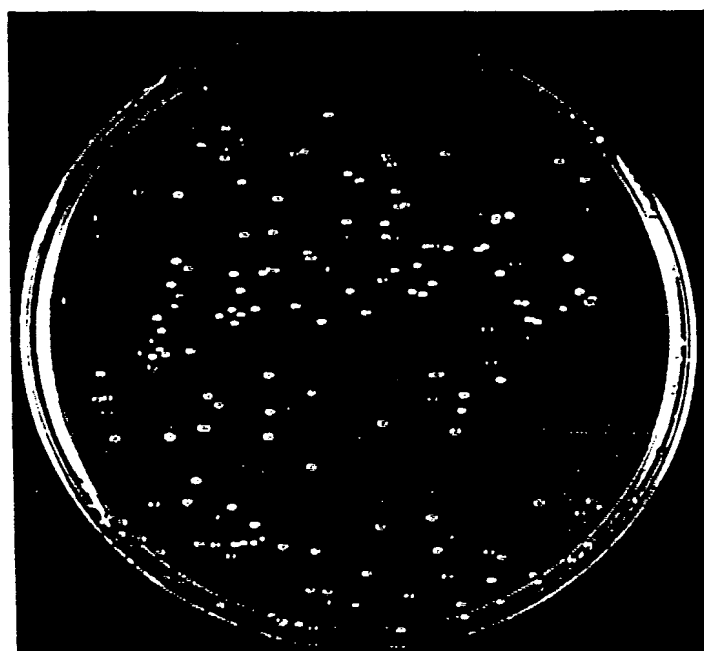


FIGURE 4B

10 Rec'd PCH/710 0 2 JUN 2020

Interacting partners

Functional complementation

β -galactosidase activity:

++ \geq 2000 U/mg
+ : 800-1500 U/mg
- \leq 300 U/mg

Polypeptide fused to the N-terminus of T18: Polypeptide fused to the C-terminus of T25:

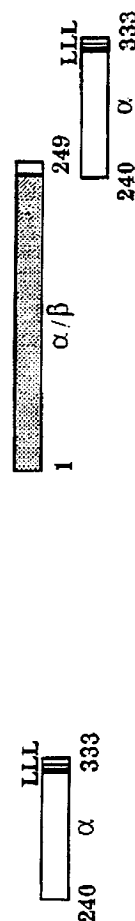
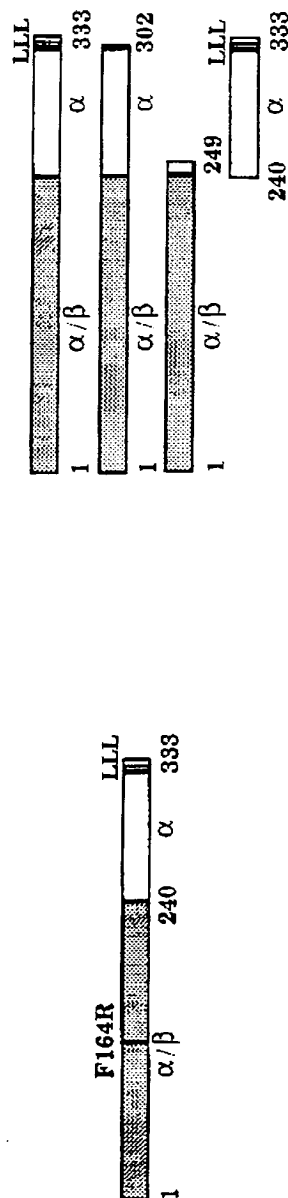
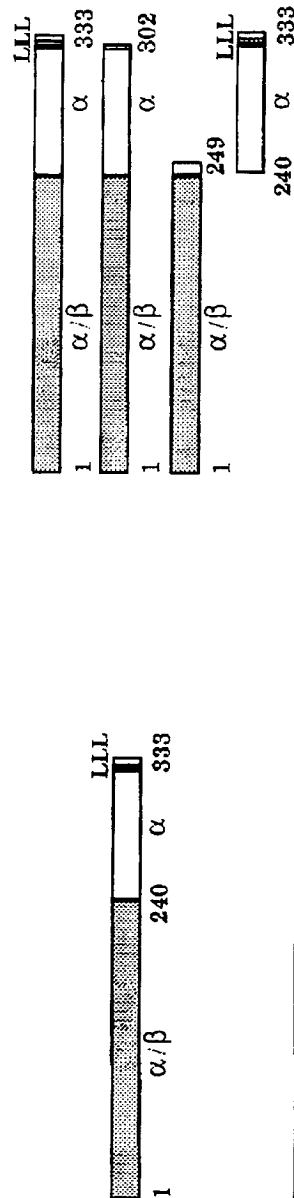


FIGURE 5

10 Rec'd FCB 11 02 JUN 2003

BvgA fragments

Functional
complementation β -galactosidase activity:

++ \geq 2000 U/mg
 + : 800-1500 U/mg
 - \leq 300 U/mg

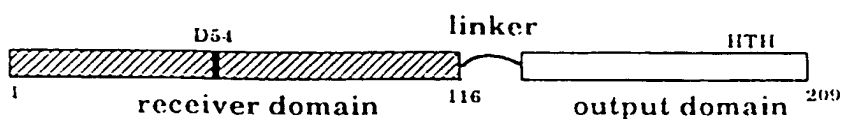
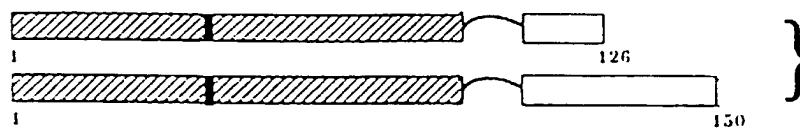
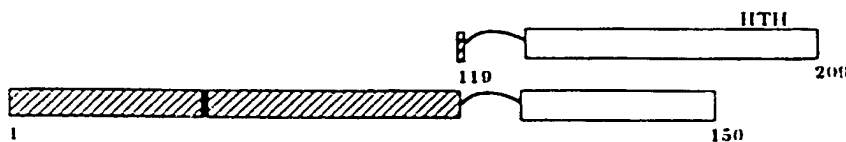
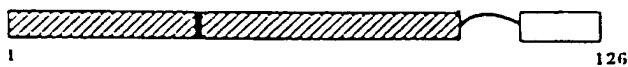
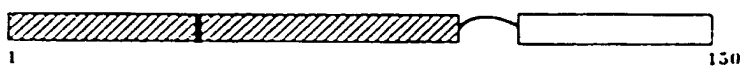
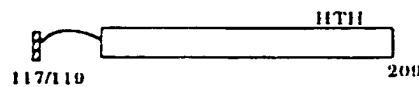
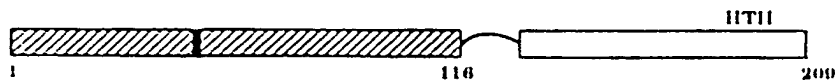
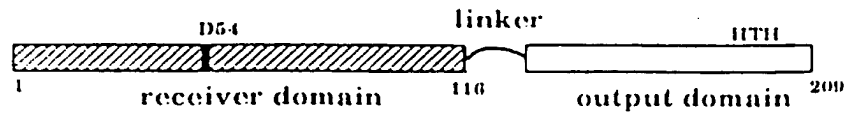


FIGURE 6

10 Rec'd FIVE

0 2 JUN 2000

A Protein Complementation Assay for Detection of Protein-Protein Interactions *in vivo*

Joelle N. Pelletier and Stephen W. Michnick, Département de biochimie, Université de Montréal, Montréal, Québec, Canada H3T 1J4.

Many processes in biology, including transcription, translation, and metabolic or signal transduction pathways, are mediated by non-covalently-associated protein complexes. Widely used methods for studying protein-protein interactions *in vivo* (such as the yeast 2-hybrid system) or *in vitro* (such as coimmunoprecipitation) are highly successful, yet present important limitations. Recently, a novel strategy for detecting protein-protein interactions *in vivo*, based on complementation of protein fragments, was demonstrated (1). Oligomerization of two proteins of interest is detected by the reassembly of a reporter enzyme, reconstituted from its complementary fragments (Fig. 1). Here we report the development of a protein complementation assay based on reconstitution of dihydrofolate reductase (DHFR) activity by complementation of defined fragments of the murine enzyme (mDHFR) in *E. coli*. The direct assay described here requires no additional endogenous factors, is used for detecting specific protein-protein interactions, and can be conveniently extended to screening cDNA libraries for protein interactions.

Design considerations: We selected mDHFR as a reporter enzyme because: (1) it is relatively small and monomeric; (2) structural and functional information about DHFR exists; (3) simple DHFR assays exist for both *in vivo* and *in vitro* measurement; and (4) overexpression of mDHFR in eukaryotic and prokaryotic cells has been demonstrated. In this study, we use GCN4 leucine zippers as model interacting proteins because their association is well characterized. DHFR is formed of three structural fragments folded as two domains: the adenine binding domain (F[2]) and a discontinuous domain (F[1] and F[3]). The folate binding pocket and the NADPH binding groove are formed mainly by residues belonging to F[1] and F[2]. We constructed mDHFR fragments F[1,2] and F[3] so as to cause minimal disruption of the active site and substrate binding sites. The native N-terminus of mDHFR and the novel N-terminus created by cleavage occur on the same surface of the enzyme facilitating N-terminal covalent attachment of each fragment to associating fragments such as the leucine zippers (Z) used here (Fig. 1).

Results: Cotransformation of *E. coli* with constructs coding for Z-F[1,2] and Z-F[3] was undertaken with the goal of detecting enzymatic activity from reconstituted "H₂" as survival of cells grown under selective pressure. Specifically, bacteria were grown in the presence of trimethoprim, which selectively inhibits bacterial DHFR.

Figure 2A illustrates that colony growth under selective pressure is possible only in cells expressing both fragments of mDHFR. The presence of the leucine zipper on both fragments of mDHFR is essential (panel II, vs III), demonstrating that oligomerization of domain Z is an absolute requirement for DHFR reassembly. Overexpressed fusion proteins of the expected molecular weight were apparent on

Figure 1: Scheme for Complementation Assays

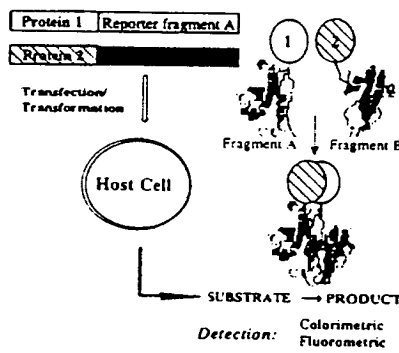
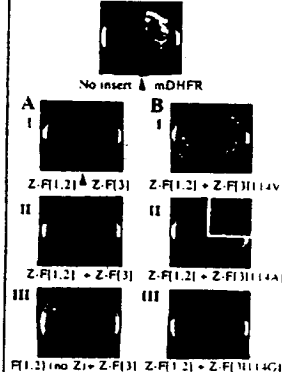


Figure 2: Survival of *E. coli*



SDS-PAGE of crude lysate.

Stability mutants: Protein stability can be reduced by changing the side-chain volume in the hydrophobic core. We generated mutants of F[3] to change the efficiency of reassembly of the mDHFR moieties into active enzyme, as tests for specific reconstitution. Thus, we mutated Ile 114 to Val, Ala, or Gly. Colonies obtained from cotransformation of Z-F[1,2] with Z-F[3]:[Ile114Ala] grew more slowly than those cotransformed with Z-F[3] or Z-F[3]:[Ile114Val] (Fig. 2B; inset to panel II is 5x enlarged). No colony growth was detected in cells cotransformed with Z-F[3]:[Ile114Gly]. The doubling times measured for cells expressing Z-F[1,2] + Z-F[3], Z-F[1,2] + Z-F[3]:[Ile114Val] and Z-F[1,2] + Z-F[3]:[Ile114Ala] were 1.6-fold, 1.9-fold and 4.1-fold, higher respectively, than the doubling time of *E. coli* expressing wild-type mDHFR.

Conclusions: We have developed a protein complementation assay based on mDHFR, where a leucine zipper directs the reconstitution of DHFR activity. Activity was detected by an *E. coli* survival assay which is both practical and inexpensive. The *in vivo* assay described here could be applied to screening cDNA libraries for the detection of unknown, specific protein-protein interactions.

References (1) N. Johnson and A. Varshavsky (1994) *PNAS* 22, 10340. (2) C. Byströff and J. Kraut (1991) *Biochemistry* 30: 2227.

THERAPEUTIC STRATEGIES FOR THE NON-INVASIVE TREATMENT OF NEURODEGENERATIVE DISEASES BY POLYAMINE-MODIFIED NEUROTROPHIC FACTORS AND ANTIOXIDANT ENZYMES WITH INCREASED PERMEABILITY AT THE BLOOD-BRAIN BARRIER AND RETAINED BIOACTIVITY.

Joseph F. Poduslo, Thomas M. Wengenack, and Geoffrey L. Curran. Molec. Neurobiol. Lab., Depart. of Neurol. and Biochem/Molec. Biol., Mayo Clinic / Foundation, Rochester, MN 55905

INTRODUCTION There is a well-characterized loss of basal forebrain cholinergic neurons associated with Alzheimer's disease (AD). Selective loss of those neurons in animal lesion and aging models produce learning and memory deficits that resemble those associated with AD. Those neurons also express large numbers of receptors for nerve growth factor (NGF) which has been shown to be necessary for the development and maintenance of these neurons. Furthermore, central administration of NGF has been shown to reduce the degeneration of those neurons in animal lesion and aging models and has been proposed for human studies (1). Other neurotrophic factors have also been isolated that support the growth of specific populations of neurons that are seen degenerate in other neurodegenerative diseases, such as amyotrophic lateral sclerosis and Parkinson's disease (1).

Recent studies have reported that free radicals may play a role in the pathogenesis of several neurodegenerative diseases, including AD, and therefore antioxidant enzymes such as superoxide dismutase (SOD) may provide another therapeutic strategy (2, 3, 4). Specifically, several studies have provided evidence that β -amyloid neurotoxicity may be mediated by the production of free radicals (5, 6, 7). In separate experiments, the application of β -amyloid was shown to increase free radical production (7) and disrupt calcium homeostasis (6). Disruption of calcium homeostasis can activate a cascade of several potentially neurodegenerative processes such as cytoskeletal breakdown and tau hyperphosphorylation, which may lead to neurofibrillary tangles (3). Furthermore, the presence of superoxide dismutase and catalase reduced the neurotoxicity of β -amyloid (5). Neurotrophic factors and antioxidant enzymes, however, have relatively low permeability at the blood-brain barrier (BBB) and must be administered directly into the brain to exhibit neuroprotective effects, particularly in the case of NGF (1). Such invasive techniques pose serious problems for clinical application of these therapies. Previous studies in our laboratory have shown that the permeability of peptides and proteins at the BBB can be significantly increased through modification with various methods (8,9,10).

POLYAMINE MODIFICATION OF NEUROTROPHIC FACTORS AND ANTIOXIDANT ENZYMES We have developed several strategies to facilitate the non-invasive, targeted delivery of therapeutic proteins with preserved biological activity across the BBB for the treatment of neurological disease (9,10). One strategy involves the covalent attachment of the naturally-occurring polyamine, putrescine (PUT), to proteins to increase their permeability at the BBB after systemic administration (10). The permeability of NGF and SOD at the BBB, before and after covalent modification with PUT, was quantified in normal adult rats using an *i.v.* bolus injection technique (3). Briefly, protein labeled with 125 I was injected rapidly into the brachial vein. Blood

was sampled during the next 30 min from the brachial artery. Another aliquot of protein, labeled with 125 I, was then injected into the brachial vein 15 sec prior to sacrifice of the animal to serve as an indicator of residual plasma volume (V_p). Brains were removed, dissected, and assayed in a two-channel gamma counter. The permeability coefficient \times surface area products (PS) for the native and modified proteins were then calculated using the V_p as a measure of residual plasma volume. The PS values of PUT-NGF at the BBB were significantly increased in six different brain regions and ranged from 4.4 to 6.2-fold greater than native NGF. The PS values of PUT-SOD at the BBB were also significantly increased in the six brain regions and ranged from 7.8 to 9.1-fold greater than native SOD.

CONCLUSIONS Polyamine modification, therefore, can dramatically increase the permeability of proteins at the BBB. There were, however, no significant changes in the V_p values for PUT-NGF or PUT-SOD indicating that PUT specifically increased permeability of the proteins without non-specific effects on the vasculature. Similar results were observed with other diverse proteins such as insulin, albumin, and IgG (10). Functionally, PUT-SOD has been shown to be neuroprotective in an animal model of global cerebral ischemia following systemic administration while native SOD was not (11). Systemic administration of polyamine modified proteins might prove to be an efficient, non-invasive approach to deliver therapeutic agents into the CNS for the treatment of AD and other neurodegenerative diseases.

REFERENCES

- Hefli F, Venero JL, Widmer HR, Knusel B. In: Iqbal K, Mortimer JA, Winblad B, Wisniewski HM (eds) Research advances in AD and related disorders. John Wiley and Sons, 1993: 321-328.
- Smith MA, Perry G. Free radical damage, iron, and AD. *J Neurol Sci* 1995; 134: 92-94.
- Mattson MP. Calcium and neuronal injury in AD. *Ann N Y Acad Sci* 1994; 747: 50-76.
- Beal MF. Aging, energy, and oxidative stress in neurodegeneration. *Ann Neurol* 1995; 38: 357-366.
- Manelli AM, Puttfarcken PS. $\text{A}\beta$ -induced toxicity in rat hippocampal cells: *in vitro* evidence for the involvement of free radicals. *Brain Res Bull* 1995; 38: 569-576.
- Mark RJ, Hensley K, Butterfield DA, Mattson MP. $\text{A}\beta$ impairs ion-motive ATPase activities. *J Neurosci* 1995; 15: 6239-6249.
- Thomas T, Thomas G, McLendon C, Sutton T, Mullan M. β -amyloid-mediated vasoactivity and vascular endothelial damage. *Nature* 1996; 380: 168-171.
- Poduslo JF, Curran GL, Berg CT. Macromolecular permeability across the blood nerve and blood-brain barriers. *Proc Natl Acad Sci USA* 1994; 91: 5705-5709.
- Poduslo JF, Curran GL. Glycation increases the permeability of proteins across the blood-nerve and blood-brain barriers. *Molec Brain Res* 1994; 23: 157-162.
- Poduslo JF, Curran GL. Polyamine modification increases the permeability of proteins at the blood-nerve and blood-brain barriers. *J Neurochem* 1996; 66: 1599-1609.
- Wengenack TM, Curran GL, Poduslo JF. Posts ischemic, systemic administration of polyamine-modified superoxide dismutase reduces hippocampal Ca^{2+} neurodegeneration in rat global cerebral ischemia. Submitted.

XP-002064565

Monitoring protein–protein interactions in intact eukaryotic cells by β -galactosidase complementation

FABIO ROSSI, CAROL A. CHARLTON, AND HELEN M. BLAU*

Department of Molecular Pharmacology, Stanford University School of Medicine, Stanford, CA 94305-5332

Communicated by Lubert Stryer, Stanford University School of Medicine, Stanford, CA, June 3, 1997 (received for review April 20, 1997)

ABSTRACT We present an approach for monitoring protein–protein interactions within intact eukaryotic cells, which should increase our understanding of the regulatory circuitry that controls the proliferation and differentiation of cells and how these processes go awry in disease states such as cancer. Chimeric proteins composed of proteins of interest fused to complementing β -galactosidase (β -gal) deletion mutants permit a novel analysis of protein complexes within cells. In this approach, the β -gal activity resulting from the forced interaction of nonfunctional weakly complementing β -gal peptides ($\Delta\alpha$ and $\Delta\omega$) serves as a measure of the extent of interaction of the non- β -gal portions of the chimeras. To test this application of *lacZ* intracistronic complementation, proteins that form a complex in the presence of rapamycin were used. These proteins, FRAP and FKBP12, were synthesized as fusion proteins with $\Delta\alpha$ and $\Delta\omega$, respectively. Enzymatic β -gal activity, served to monitor the formation of the rapamycin-induced chimeric FRAP/FKBP12 protein complex in a time- and dose-dependent manner, as assessed by histochemical, biochemical, and fluorescence-activated cell sorting assays. This approach may prove to be a valuable adjunct to *in vitro* immunoprecipitation and crosslinking methods and *in vivo* yeast two-hybrid and fluorescence energy transfer systems. It may also allow a direct assessment of specific protein dimerization interactions in a biologically relevant context, localized in the cell compartments in which they occur, and in the milieu of competing proteins.

Specific interactions between proteins in mammalian cells are the basis of many essential biological processes. For example, protein–protein interactions are involved in the assembly of enzymes and other protein homodimers and heterodimers that play important roles in the regulation of intracellular transport pathways, gene expression, receptor–ligand interactions, and in the therapeutic or toxic effects of administered drugs. To increase our understanding of these biological processes, several techniques have been developed for examining the interactions between proteins within cells. Coimmunoprecipitation experiments with antibodies are suggestive of such interactions in that they allow a determination of the affinity of a given protein for another protein, albeit *in vitro* following cell lysis under conditions that cannot determine whether the two proteins are present within the same compartment or at the concentrations tested (1–3). Methods for crosslinking proteins within the cell and then cofractionating them by chromatography have also proven useful, although purification, sequencing, and identification of the crosslinked proteins can be difficult when they are present in small quantities.

In addition to such biochemical techniques, the yeast two-hybrid system has been extremely useful for detecting and identifying protein–protein interactions *in vivo* (4–6). This system takes advantage of the properties of the GAL4 protein of the

yeast *Saccharomyces cerevisiae*, a transcriptional activator required for the expression of genes encoding enzymes involved in galactose utilization. The GAL4 protein consists of two separate and identifiable domains, an N-terminal DNA binding domain and a C-terminal transcription activation domain. Separate fusion proteins, each comprising only one of the two GAL4 domains fused to one of two different test polypeptides, interact through affinity of the different test polypeptides, bringing two GAL4 domains into close physical proximity and reconstituting GAL4 function. A distinct advantage of this approach over biochemical approaches is that it allows identification of novel protein partners at a molecular level. However, the system requires that protein–protein interactions occur in the nucleus of a cell leading to transcriptional activation of a reporter gene and the detection of a diffusible product. Thus, the assay is indirect and is dependent on other cellular functions. Nonetheless, numerous previously unknown protein interactions have been identified using the yeast two-hybrid system.

Fluorescence ratio imaging has also been used to study protein interactions in live cells (7). This innovative system has yielded several important new findings. However, it is limited by the requirement that the fluorescent labels on the interacting proteins be sufficiently close to permit efficient energy transfer. Also, the labeled proteins need to be introduced into the cells at relatively high concentrations. Clearly, a method that would allow a direct examination of molecular interactions with fewer size constraints, at the site where they occur within a eukaryotic cell, would be advantageous.

Here we describe a novel application of the bacterial *lacZ* gene that may allow the direct detection of protein–protein interactions *in situ* in a range of cell types and species. The product of the *lacZ* gene, β -galactosidase (β -gal), has been used for many years as a reporter gene to measure transcriptional activity by histochemical or biochemical assays or by live cell sorting (8–11). A property of the *lacZ* gene, intracistronic complementation, has been known and studied for many years in prokaryotes (12–14), but has only recently been adapted for use in eukaryotes (15). Pairs of inactive β -gal deletion mutants are capable of complementing one another in trans and assembling to form an active enzyme. Pairs of deletion mutants that complement with either high or low efficiency were identified (15, 16). In previous studies of myoblast fusion, the most efficient complementing pair was employed to analyze the putative role of cell adhesion and signal transduction molecules in this poorly understood process (17).

As reported here, pairs of inactive β -gal deletion mutants were used to produce chimeric proteins to test whether the complemented enzyme could serve as a marker of other protein–protein interactions (see Fig. 1). Several features were critical in the application of this approach. Most important was the necessity to select β -gal mutants with sufficiently low affinity for each other so that they monitored rather than drove

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

© 1997 by The National Academy of Sciences 0027-8424/97/948405-06\$2.00/0
PNAS is available online at <http://www.pnas.org>.

Abbreviations: β -gal, β -galactosidase; FACS, fluorescence-activated cell sorting; FKBP12, FK506-binding protein-12; FRAP, FKBP-rapamycin associated protein.

*To whom reprint requests should be addressed.

the association of the test proteins. Our biochemical, histochemical, and fluorescence-activated cell sorting (FACS) experiments show that appropriate β -gal mutants can be incorporated into fusion proteins that serve as tracers of protein-protein interactions in intact eukaryotic cells.

MATERIALS AND METHODS

Construction of pWZL FRAP- $\Delta\alpha$ and pWZL FKBP12- $\Delta\omega$ Viruses. To be able to fuse the sequences coding for FKBP12 and the FKBP12-rapamycin binding domain in-frame with β -gal, an adaptor oligonucleotide (CATGGAGCTCGAGAG) containing an *Xho*I site was inserted in the *Nco*I site at the ATG of the previously described $\Delta\alpha$ and $\Delta\omega$ β -gal mutants (15). Two *Xho*I-*Sal*I DNA fragments corresponding to amino acids 2025–2114 of human FRAP and to the complete coding sequence of human FKBP12 (kind gifts of Gerald Crabtree, Stanford University School of Medicine) were cloned into the *Xho*I site of the modified $\Delta\alpha$ and $\Delta\omega$ mutants, generating FRAP- $\Delta\alpha$ and FKBP12- $\Delta\omega$. For both constructs, conservation of the appropriate reading frame was confirmed by sequencing.

To insert the FRAP- $\Delta\alpha$ and FKBP12- $\Delta\omega$ coding sequences in pWZL-neo and pWZL-hygro retroviruses (P. J. Morgenstern, unpublished work), an adaptor oligonucleotide containing *Nco*I and *Bam*HI sites (GATCACCATGGACGCGTGGATCCC) was inserted in the *Bam*HI and *Xho*I sites of the pWZL vectors. Both of the original sites were destroyed by this insertion. The FRAP- $\Delta\alpha$ and FKBP12- $\Delta\omega$ coding sequences were then inserted in the modified pWZL vectors as *Nco*I-*Bam*HI fragments.

Virus Production and Infection Protocol. Proviral constructs were introduced into the Phoenix cotropic packaging cells (P. L. Achacoso and G. P. Nolan, unpublished work) by calcium phosphate transfection. The media-containing retrovirus from the packaging cells was harvested 24–72 hr after transfection and used to infect C2C12 myoblasts (18) in the presence of 8 μ g/ml polybrene (Sigma). Singly and doubly infected cells were selected with the appropriate drugs. Both G418 and hygromycin were used at a final concentration of 1 mg/ml. The selected cells were expanded as populations for subsequent experiments.

Quantitation of β -Gal Activity. Unless otherwise stated, rapamycin (Calbiochem) was used at a concentration of 10 ng/ml.

Biochemical quantitation of β -gal activity. β -gal activity was measured by chemiluminescence as described (15). Briefly, cells cultured in microtiter plates were lysed *in situ* in 50 μ l of a 1:1 mixture of lysis and assay buffers containing Galacton Plus substrate from the Galacto-Light Plus assay kit (Tropix, Bedford, MA). Reactions were terminated after 1 hr at room temperature. After addition of Light Emission Accelerator solution, luminescence was measured using a MicroBeta 1450 scintillation counter (Wallac, Gaithersburg, MD).

Histochemical detection of β -gal using fluorescence histochemistry (Fluor-X-Gal). Cells were processed as described (15). Briefly, cells grown on glass coverslips were fixed in 4% paraformaldehyde in PBS and rinsed twice with PBS. For triple-labeling (actin, β -gal, and nuclei), fixed cells were first stained for actin by incubation in a solution of 1 μ M biotin-XX-phalloidin (Molecular Probes) for 30 min at room temperature, followed by PBS rinses and incubation in Cy5-labeled streptavidin (Amersham) (1:250 in PBS, 30 min, room temperature). β -gal was detected by incubation of fixed cells in a solution of 25 μ g/ml 5-bromo-6-chloro-3-indolyl β -D galactopyranoside (Fluka) plus 100 μ g/ml fast red violet LB (Sigma) in PBS for 45 min at 37°C. Coverslips were rinsed 4 times with PBS, nuclei labeled with Hoechst 33258 (Calbiochem; 1:10,000 in PBS), rinsed again, mounted in PBS, and sealed to glass slides. Triple-labeled deconvolved images were collected using a DeltaVision deconvolution microscope (Applied Precision, Mercer Island, WA). Images were represented in false color using Adobe Photoshop software. Fluor-X-Gal staining can be detected with either fluorescein or rhodamine filters; in this case, detection was with the rhodamine filter, and is depicted

in green (see Fig. 3 C and D). Cy5 was detected with a Cy5 filter and is depicted in red; Hoechst was detected with a DAPI filter and is depicted in blue (see Fig. 3 C and D). Double-labeled cells were photographed using a Zeiss Axiophot fluorescence microscope. Fluor-X-Gal staining was detected with a rhodamine filter and appears red; Hoechst was detected with a UV filter (see Fig. 3 A and B).

FACS. β -gal expression in live cells was determined on a FACS as described (10), except that the fluorescein di- β -D-galactopyranoside substrate was used at a concentration of 1 mM.

RESULTS

Design of Fusion Protein Test System for Monitoring Interactions Using Complementing β -Gal Mutants. We have recently adapted *lacZ* intracistronic complementation for use in eukaryotic cells (15). For the purpose of testing the potential application of β -gal complementation as a method to monitor protein-protein interactions, a pair of mutants was selected that contained the domains postulated to be necessary for trans-complementation, but which were impaired in their ability to restore enzymatic activity upon coexpression in mammalian cells. These mutants were the classical ω donor M15 (hereafter referred to as $\Delta\alpha$), which lacks amino acids 11–41 of the wild-type molecule, and an unusually long α donor containing the first 788 amino acids of β -gal ($\Delta\omega$) (15). Since both polypeptides are capable of efficiently complementing a third deletion mutant ($\Delta\mu$, lacking 553 amino acids in the central portion of the molecule), it seemed unlikely that their poor ability to complement each other was due to instability or to low expression. Inefficient α donors have long been known to exist in prokaryotes, and differences in their folded structure that render the α domain unavailable for complementation have been proposed as a basis for their biochemical properties (16). One model separates the β -gal complementation mechanism into two steps: rapid formation of a complex between α and ω donor peptides, followed by a slow conformational change from an inactive to an active conformation (19, 20). Which of these two steps was impaired in inefficient α donors was unclear. We reasoned that if complex formation but not catalytic activity were impaired, it should be possible to obtain efficient complementation by promoting the heterodimerization of weakly complementing $\Delta\alpha$ and $\Delta\omega$ β -gal mutants.

To test this hypothesis, we used a well-characterized protein complex (21–24). The intracellular rapamycin binding protein, FK506-binding protein-12 (FKBP12), interacts with the intracellular FKBP-rapamycin associated protein (FRAP) only when rapamycin is present in the culture medium. This interaction is well documented to increase over time and with the dose of rapamycin. Rapamycin is a small, cell-permeable molecule that binds to the two intracellular proteins via independent determinants. Since rapamycin is unable to bind two FKBP12 molecules at the same time and FRAP only binds rapamycin within the FKBP12-rapamycin complex, heterodimers do not form unless rapamycin is present (25).

Specifically, the test system involved combining the weakly complementing β -gal mutants, $\Delta\alpha$ or $\Delta\omega$, with the FKBP12/FRAP/rapamycin system by producing fusion proteins. Two different retroviral constructs were designed that encoded fusion proteins of either the FKBP12-rapamycin binding domain of FRAP or the entire FKBP12 peptide together with the β -gal $\Delta\alpha$ or $\Delta\omega$ mutants, respectively (FRAP- $\Delta\alpha$ and FKBP12- $\Delta\omega$). A prediction of the hypothesis is that introduction of these fusion proteins into cells would result in negligible β -gal activity in the absence of rapamycin (Fig. 1A). Thus, β -gal activity would first become detectable upon addition of drug. As a result, β -gal activity would monitor the association of FRAP and FKBP12 proteins (green and red proteins and cDNAs in Fig. 1 B and C, respectively).

A critical feature of the design of the system presented here was to reduce protein expression levels as much as possible to avoid perturbing the intracellular protein milieu. Accordingly,

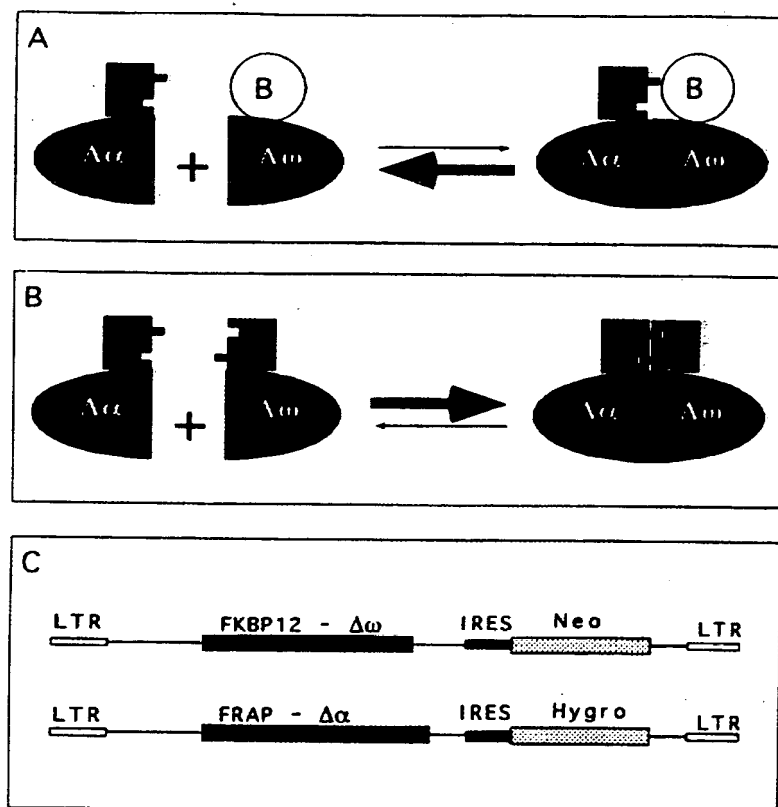


FIG. 1. Experimental design. (A) When the $\Delta\alpha$ and $\Delta\omega$ β -gal mutants are fused to proteins that do not dimerize, their association is not favored and β -gal activity is not detected. (B) When the $\Delta\alpha$ and $\Delta\omega$ β -gal mutants are fused to proteins that can dimerize, the formation of active β -gal is favored. (C) Schematic representation of the FKBP12- $\Delta\omega$ -Neo and the FRAP- $\Delta\alpha$ -Hygro constructs. IRES, internal ribosome entry sequence; LTR, long terminal repeat.

the fusion proteins should serve as "tracers" of naturally occurring protein-protein interactions, overcoming the problems previously experienced upon overexpression of introduced proteins either by transient or stable transfection, or transduction with retroviruses. To reduce expression levels, in the experiments reported here the cDNAs encoding FRAP- $\Delta\alpha$ and FKBP12- $\Delta\omega$ were not inserted into MFG, but into the pWZL retroviral vector (P. J. Morgenstern, unpublished work), upstream of the gene encoding hygromycin or neomycin, respectively. The levels of protein expression are reduced in the pWZL vector due to the presence of mutations that delete the splice donor/acceptor sequences upstream of the ATG of the fusion proteins. These mutations result in a lower translation efficiency of the first coding sequence of a bicistronic message, but do not affect the translation of the second sequence, in this case the selectable marker, which is solely dependent on an encephalomyocarditis virus internal ribosomal entry sequence. Using a pWZL vector, 50% less of the upstream protein is expressed compared with vectors containing wild-type splice donor/acceptor sequences (data not shown). As a result of the reduced levels of expression the frequency of spontaneous interactions of β -gal mutants, which is concentration dependent, should be significantly reduced.

Induction of β -Gal Activity on Coexpression of FRAP- $\Delta\alpha$ and FKBP12- $\Delta\omega$ Fusion Proteins in the Presence of Rapamycin. The FRAP- $\Delta\alpha$ -Neo and FKBP12- $\Delta\omega$ -Hygro vectors were tested in an established line of myoblasts, C2C12 (18). Infectious viral particles were produced by transient transfection of each construct into the Phoenix packaging cell line (P. L. Achacoso and G. P. Nolan, unpublished work). C2C12 myoblasts were infected either singly with each retrovirus alone or simultaneously with both. All experiments were performed after selection with hygromycin and G418 to ensure that 100% of the cells contained the constructs.

β -gal activity was measured in biochemical assays and found to be dependent on the presence of rapamycin in the medium. C2C12 cells expressing both fusion proteins were plated in

replicate in 96-well plates. Rapamycin was added to the culture medium, and the β -gal activity was measured at different time points. For each time point, six replicate samples were assayed with a sensitive chemiluminescence assay, as described (15). In untreated control samples, no β -gal activity was detected above background. Rapamycin at a concentration of 10 ng/ml induced a 30-fold increase in β -gal activity within 5 hr. After 5 hr, β -gal activity continued to increase, reaching a level 2 orders of magnitude above background within 20 hr (Fig. 24). In control populations of cells expressing only one of the two constructs, β -gal activity did not increase above background when rapamycin was added (data not shown). The linear increase in β -gal activity observed between 5 and 20 hr after rapamycin addition may be due to an increase in protein concentration resulting from cell proliferation, since the C2C12 cells have a doubling time of 12 hr. It is also possible that α donor peptides are stabilized when they are incorporated into a multimeric complex. In support of the latter possibility, Western blotting of cellular extracts with antibodies to β -gal revealed an increase in steady-state levels of β -gal mutant peptides upon rapamycin treatment (data not shown); in addition, stabilization of complementing β -gal peptides upon assembly of active multimeric enzymes has been reported by others (26).

In Fig. 2B, the dose response curve is shown. β -gal activity increased linearly with the dose of rapamycin in the 0–10 ng/ml range. This linearity suggests that β -gal enzymatic activity can serve as a reporter to quantitate protein-protein interactions. The results of these experiments demonstrate that the interaction of the FKBP12- and FRAP- β -gal fusion proteins in the presence of rapamycin is specific and exhibits a comparable dose-response curve to results previously obtained by others for the FKBP12/FRAP/rapamycin complex alone (25). Thus, fusion to β -gal peptides does not interfere with the interaction of the FKBP12 and FRAP proteins. Moreover, endogenous FKBP12 and FRAP proteins are ubiquitously expressed and will interact in the presence of rapamycin, thereby competing with the introduced FRAP- $\Delta\alpha$ and FKBP12- $\Delta\omega$ fusion proteins, yet not generating

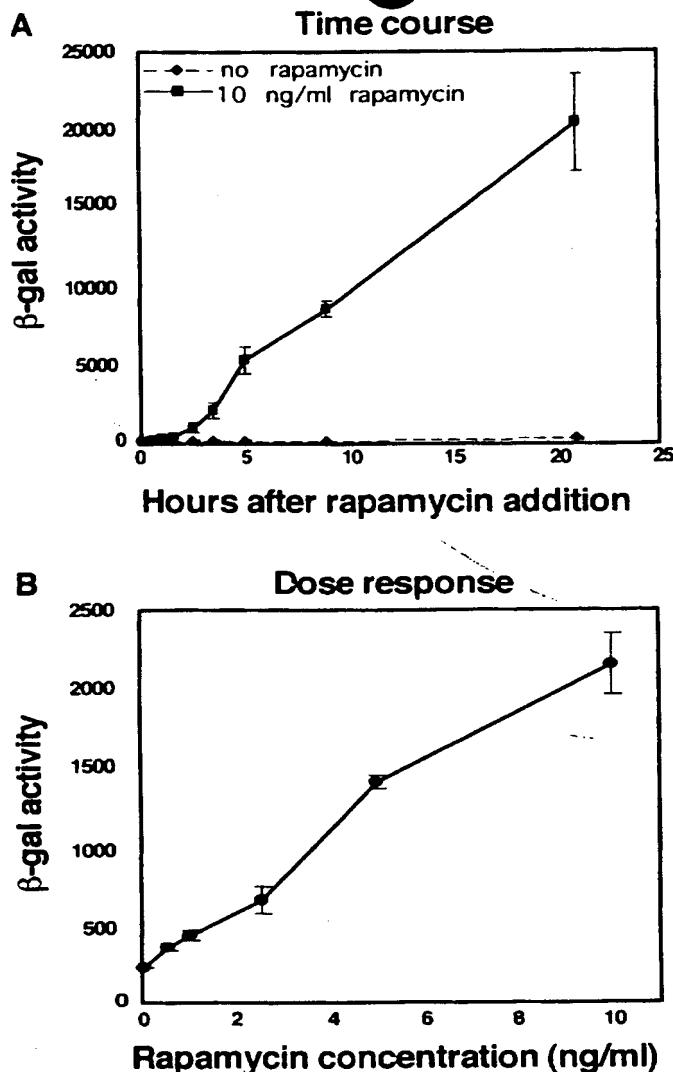


Fig. 2. Biochemical assay of induction of β -gal activity upon chimera complementation. (A) Kinetics of induction of β -gal activity upon treatment with rapamycin. Pure populations of C2C12 cells stably expressing both FKBP12- $\Delta\omega$ and FRAP- $\Delta\alpha$ were plated in 96-well plates and 10 ng/ml rapamycin was added at time zero. Cells were then lysed at different time intervals thereafter, and the β -gal activity in the lysates was quantitated by chemiluminescence. (B) Dose response of β -gal activity upon rapamycin treatment. C2C12 cells expressing both FKBP12- $\Delta\omega$ and FRAP- $\Delta\alpha$ were plated in 96-well plates and treated with different concentrations of rapamycin for 3.5 hr. β -gal activity is expressed as luminescence counts per second. Each point represents the average of six replicate samples. Error bars indicate standard deviations from the mean.

β -gal activity. Our results indicate that productive FRAP- $\Delta\alpha$ and FKBP12- $\Delta\omega$ dimers will also form, generating β -gal activity that will reflect the interaction of FRAP with FKBP12-rapamycin in that cellular environment even in the presence of the competing endogenous proteins. This finding suggests that it will be possible to use β -gal complementation as a tool to analyze protein-protein interactions generally.

β -gal activity was assayed using the sensitive Fluor-X-Gal histochemical stain, which allows simultaneous tricolor fluorescence analysis, as described (15). Because of its broad emission spectrum, the fluorescence of this substrate can be visualized at wavelengths that yield either a red (Fig. 3A and B) or a green signal (Fig. 3C and D). In the absence of rapamycin, only a very weak β -gal fluorescence was observed. For example, in Fig. 3A,

in which two fluorochromes were used at low magnification, the blue nuclei of the myoblasts was primarily evident with little evidence of red β -gal activity. Similarly, in Fig. 3C, in which three fluorochromes were used at higher magnification, blue nuclei and red phalloidin-staining of actin were evident, but green β -gal activity was barely detectable. In both histochemical assays, in the presence of 10 ng/ml rapamycin, the activity of β -gal increased significantly to produce an intense stain that was primarily localized in the cytoplasm, as expected (Fig. 3B and D). Detection of the complemented β -gal enzyme was enhanced by imaging with a deconvolution microscope (Fig. 3D). These images show that protein complexes forming outside the nucleus are readily detectable using this assay.

The β -gal activity of a population of cells was assayed in the presence and absence of 10 ng/ml rapamycin by FACS. Using this sensitive assay, we were able to detect increased β -gal activity in most of the cells after only 90 min of rapamycin treatment (Fig. 4A). A range of expression levels was seen, as evidenced by the breadth of the peak of emission in the presence and absence of the drug (Fig. 4A, blue and red profiles). This breadth is presumably due to variable efficiency of expression of each of the retroviral vectors following integration in the target cell. This inference is supported by the finding that when the 25% of the cells expressing the lowest β -gal activity in the absence of rapamycin were collected (Fig. 4B) and then reassayed in the presence and absence of rapamycin, a clear distinction between the two populations was seen (Fig. 4C). Thus, nonoverlapping populations of cells that do or do not express complementing fusion proteins can be identified and is slated by FACS.

Rapamycin-Dependent Induction of β -Gal Activity in Cell Lysates. To test whether the heterodimerization of FRAP- $\Delta\alpha$ and FKBP12- $\Delta\omega$ induced by rapamycin required cotranslational assembly or could occur with fully synthesized and folded proteins, β -gal activity was assayed in a cell free system. Cells expressing both constructs were grown in the absence of rapamycin and lysed *in situ* in 50 μ l of a 1:1 mixture of lysis and assay buffers containing Galacton Plus substrate from the Galacto-Light Plus assay kit (Tropix). Rapamycin was then added to the lysates, and β -gal activity was quantitated immediately, 1 and 3 hr later. As a control, β -gal activity was measured in a parallel set of lysates that were not exposed to rapamycin at all. No statistically significant increase in β -gal activity was detected in the samples that did not receive rapamycin. By contrast, a more than 2-fold increase in β -gal activity was observed in the rapamycin-treated lysates 1 hr after drug administration. The increase in activity detected in lysates is only a fraction of the increase observed upon rapamycin treatment of intact cells, a finding that very likely reflects the lower concentration of FRAP- $\Delta\alpha$ and FKBP12- $\Delta\omega$ in the lysates. This is probably due to the combined effects of the instability of β -gal mutants within lysates and of their dilution during lysis. Nevertheless, these experiments demonstrate that *de novo* synthesis is not required for complementation and that folded proteins can be induced to form complexes that can be monitored by β -gal activity.

DISCUSSION

Prerequisites for Monitoring Protein-Protein Interactions by β -Gal Complementation. We have shown that β -gal activity can be used to monitor the interaction of chimeric proteins. Critical to the success of this system was the choice of two poorly complementing β -gal mutants, since strongly complementing mutants spontaneously assemble and produce functional β -gal activity detectable in the absence of any other protein constituents (15). By contrast, the weakly interacting $\Delta\alpha$ or $\Delta\omega$ β -gal mutants expressed from the pWZL vector do not yield detectable enzymatic activity, unless their local concentration is increased by forcing them to heterodimerize. This is only achieved by synthesizing them as fusion proteins in which the non- β -gal portions of the chimeras have sufficient affinity to drive the interaction.

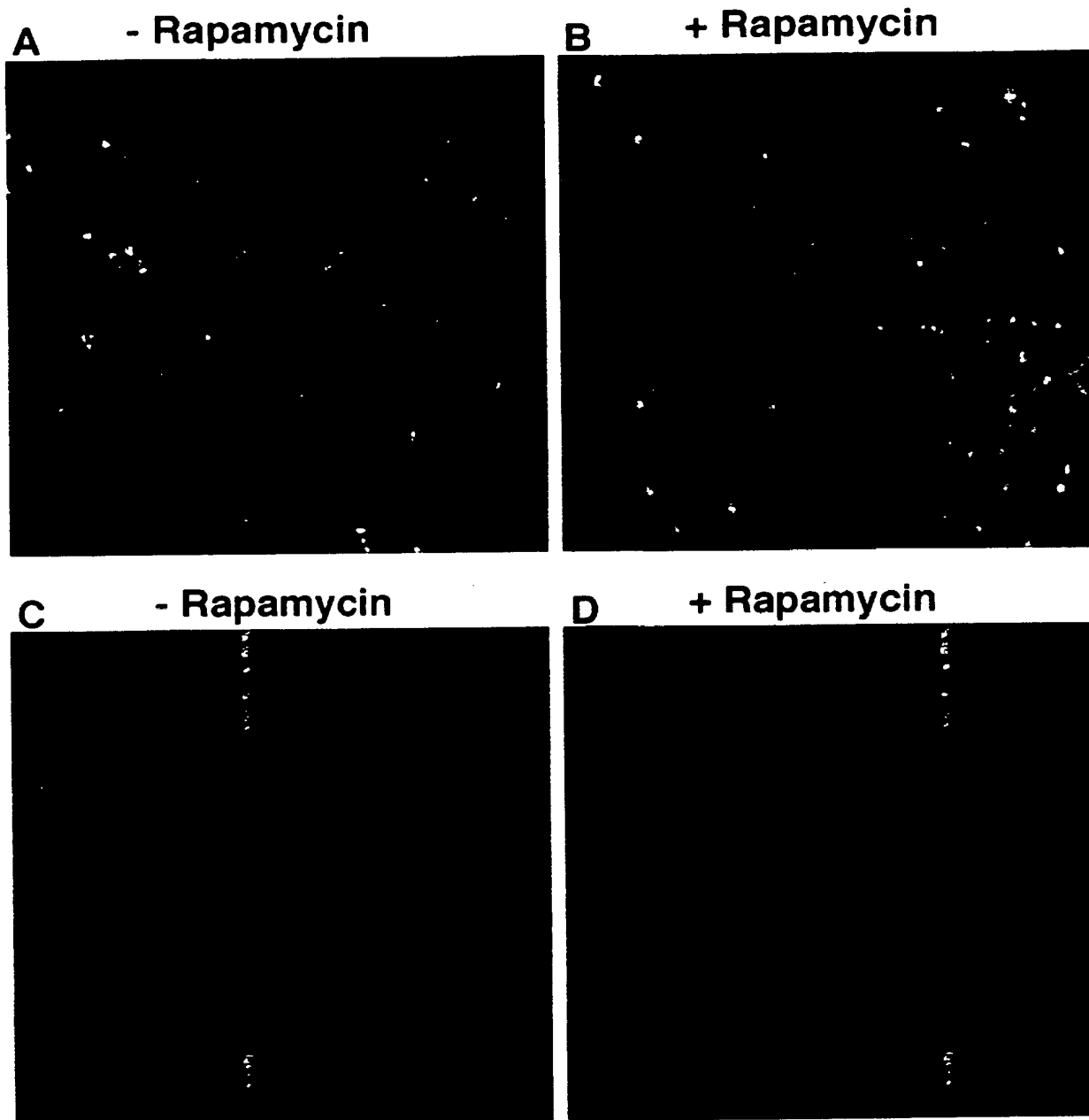


FIG. 3. Histochemical assay of induction of β -gal activity upon chimera complementation. C2C12 cells expressing both FKBP12- $\Delta\omega$ and FRAP- $\Delta\alpha$ were maintained overnight either in the absence (A and C) or the presence (B and D) of 10 ng/ml rapamycin. β -gal activity was visualized by fluorescence microscopy using Fluor-X-Gal as substrate. (A and B) Double-labeled samples showing Hoechst stained nuclei (blue) and β -gal activity using Fluor-X-Gal as substrate viewed with a rhodamine filter set (red). (C and D) Triple-labeled samples obtained by imaging with a DeltaVision microscope showing β -gal activity using Fluor-X-Gal as substrate (green), Hoechst stained nuclei (blue), and Cy5-labeled actin filaments (red) to visualize the contour of each cell at higher magnification.

The molecular basis of the impaired ability of the $\Delta\alpha$ and $\Delta\omega$ mutants to associate spontaneously and recreate active β -gal is unclear. Potential insights derive from prokaryotic studies and the recently published crystal structure of intact β -gal (27). It is well known that efficient β -gal complementation in prokaryotes requires intermolecular interactions resulting in the sharing of domains between α acceptor and ω donor mutants (20). Moreover, the recently published structure of wild-type β -gal suggests that both the α and ω domains are involved in a dimerization step critical to the assembly of functional enzymes. These domains are also involved in essential contacts with the relatively large central region of the molecule (27). $\Delta\alpha$ β -gal can be efficiently and spontaneously complemented in mammalian cells by an α donor

lacking the central μ portion of the molecule (15). Similarly, $\Delta\omega$ β -gal can complement spontaneously with an ω donor lacking the central domain. In contrast, the weakly complementing mutant pairs used here each contain a large intact central μ domain. Thus, one possible explanation for the observed lack of spontaneous complementation between the $\Delta\omega$ and the $\Delta\alpha$ peptides is that the presence of the μ domain in both mutants may diminish their affinity for each other by steric hindrance or by sequestering either the α or the ω domain in an intramolecular interaction. Accordingly, we postulate that the lack of spontaneous assembly and generation of β -gal activity can be overcome by increasing the local concentration of $\Delta\alpha$ and $\Delta\omega$ by forced association of these two mutants in chimeras, thereby counteracting the potential

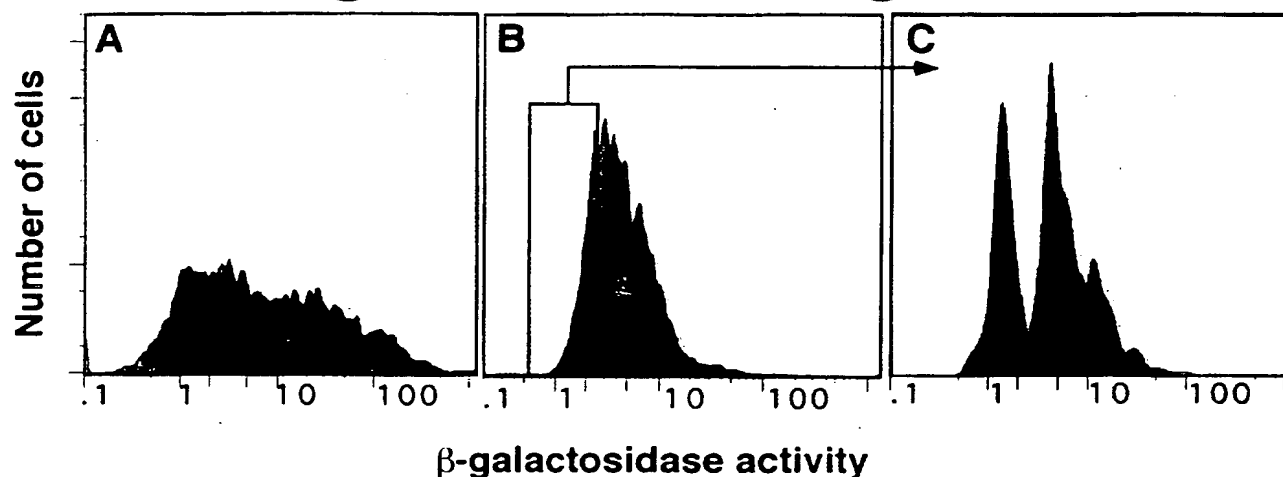


FIG. 4. FACS analysis of induced β -gal activity upon chimera complementation. The red peaks represent the untreated samples and the blue peaks represent samples treated with 10 ng/ml rapamycin. (A) Induction of β -gal activity in a population of C2C12 cells expressing both FKBP12- $\Delta\omega$ and FRAP- $\Delta\alpha$ after 90 min of rapamycin treatment. The majority of the cells respond to rapamycin treatment with an increase in β -gal activity. (B) Subpopulation of cells selected on the basis of low β -gal activity in uninduced conditions. (C) The same population was maintained overnight in the absence (red peak) or in the presence (blue peak) of rapamycin. The induced and uninduced populations yield essentially nonoverlapping peaks. The vertical axis represents relative cell number and the horizontal axis represents intensity of β -gal fluorescence on a logarithmic scale.

inhibitory effect imposed by the presence of two central domains. Testing of this model will require further mutagenesis, activity assays, and x-ray crystallographic analysis of the resulting structures.

Irrespective of the precise molecular nature of the interactions, the data presented here provide strong evidence that by engineering constructs in which domains or proteins of interest drive the dimerization of $\Delta\alpha$ or $\Delta\omega$ β -gal mutants, it will be possible to monitor such interactions by measuring the level of β -gal activity following coexpression of these fusion proteins in intact cells. Moreover, the present system should theoretically allow detection of complexes in subcellular compartments, including the nucleus, the cytoplasm, or the membrane. Finally, as shown here, protein dimerization can be monitored in the context of the cell in the presence of endogenous competing protein partners.

Potential for the Development of a "Mammalian Two-Hybrid" System. The experiments described in this report show that two distinct β -gal mutants that do not readily assemble can be forced to interact and yield significantly increased levels of active enzyme. This is achieved by coupling each mutant with one of a pair of highly interactive proteins, as chimeras or fusion proteins. In this case, we have used FRAP and FKBP12, proteins that only interact in the presence of a small molecule, rapamycin. Using this tripartite complex, we were able to show that the affinity of the non- β -gal proteins drove the interaction and that the β -gal components generated enzymatic activity as a result, serving to monitor the extent of that interaction. The levels of β -gal activity in the presence and absence of forced dimerization were clearly separable by both biochemical and FACS assays, suggesting that this system could be used to screen for as yet unidentified protein partners. The target protein fused to a complementing β -gal mutant (bait) could be stably expressed in a well-characterized cell line. Expression libraries containing cDNAs fused to a β -gal deletion mutant could be introduced into these cells with high efficiency using retroviral vectors (28). Finally, gene products that interact with the bait could be isolated by identifying β -gal positive clones. A potential advantage of this system over systems described previously is that the screen could be carried out in any cell type with its own particular milieu of competing resident proteins. An attractive possibility is that the bait could be targeted to a given cellular compartment, with the aim of identifying proteins involved in interactions restricted to that specific location. This mammalian "two-hybrid" screen could also be carried

out in the presence of extracellular signaling molecules, growth factors, or differentiation factors, that might alter the potential for dimerization of two given proteins in particular cell types. Taken together, this approach to the study of protein-protein interactions should greatly enhance our understanding of the development of diverse cell types and organisms and how that development goes awry.

We are grateful to Bruce Blakely for constructive comments on the manuscript, to Oivind Guicherit and Norris Turner for technical advice, and to Gerald Crabtree for useful discussions. F.R. was supported by a long-term postdoctoral fellowship from the Human Frontiers in Science Program. H.M.B. is the recipient of an National Institutes of Health MERIT award.

- Gu, W., Schneider, J. W., Condorelli, G., Kaushal, S., Mahdavi, V. & Nadal-Ginard, B. (1993) *Cell* 72, 309–324.
- Haley, O., Novitsch, B. G., Spicer, D. B., Skapek, S. X., Rhee, J., Hannon, G. J., Beach, D. & Lassar, A. B. (1995) *Science* 267, 1018–1021.
- Schneider, J. W., Gu, W., Zhu, L., Mahdavi, V. & Nadal-Ginard, B. (1994) *Science* 264, 1467–1471.
- Bai, C. & Elledge, S. J. (1996) *Methods Enzymol.* 273, 331–347.
- Fields, S. & Song, O. (1989) *Nature (London)* 340, 245–246.
- Chien, C. T., Bartel, P. L., Sternglanz, R. & Fields, S. (1991) *Proc. Natl. Acad. Sci. USA* 88, 9578–9582.
- Adams, S. R., Harootyan, A. T., Buechler, Y. J., Taylor, S. S. & Tsien, R. Y. (1991) *Nature (London)* 349, 694–697.
- Lis, J. T., Simon, J. A. & Sutton, C. A. (1983) *Cell* 35, 403–410.
- Nielsen, D. A., Chou, J., MacKrell, A. J., Casadaban, M. J. & Steiner, D. F. (1983) *Proc. Natl. Acad. Sci. USA* 80, 5198–5202.
- Nolan, G. P., Fiering, S., Nicolas, J. F. & Herzenberg, L. A. (1988) *Proc. Natl. Acad. Sci. USA* 85, 2603–2607.
- Price, J., Turner, D. & Cepko, C. (1987) *Proc. Natl. Acad. Sci. USA* 84, 156–160.
- Ullmann, A., Jacob, F. & Monod, J. (1967) *J. Mol. Biol.* 24, 339–343.
- Ullmann, A., Jacob, F. & Monod, J. (1968) *J. Mol. Biol.* 32, 1–13.
- Ullmann, A., Perrin, D., Jacob, F. & Monod, J. (1965) *J. Mol. Biol.* 12, 918–923.
- Mohler, W. A. & Blau, H. M. (1996) *Proc. Natl. Acad. Sci. USA* 93, 12423–12427.
- Villarejo, M., Zamenhof, P. J. & Zabin, I. (1972) *J. Biol. Chem.* 247, 2212–2216.
- Charlton, C. A., Mohler, W. A., Radice, G. L., Hynes, R. O. & Blau, H. M. (1997) *J. Cell Biol.*, in press.
- Blau, H. M., Chiu, C. P. & Webster, C. (1983) *Cell* 32, 1171–1180.
- Langley, K. E. & Zabin, I. (1976) *Biochemistry* 15, 4866–4875.
- Zabin, I. (1982) *Mol. Cell. Biochem.* 49, 87–96.
- Brown, E. J., Albers, M. W., Shin, T. B., Ichikawa, K., Keith, C. T., Lane, W. S. & Schreiber, S. L. (1994) *Nature (London)* 369, 756–758.
- Chen, J., Zheng, X. F., Brown, E. J. & Schreiber, S. L. (1995) *Proc. Natl. Acad. Sci. USA* 92, 4947–4951.
- Belshaw, P. J., Ho, S. N., Crabtree, G. R. & Schreiber, S. L. (1996) *Proc. Natl. Acad. Sci. USA* 93, 4604–4607.
- Choi, J., Chen, J., Schreiber, S. L. & Clardy, J. (1996) *Science* 273, 239–242.
- Ho, S. N., Biggar, S. R., Spencer, D. M., Schreiber, S. L. & Crabtree, G. R. (1996) *Nature (London)* 382, 822–826.
- Lin, S. & Zabin, I. (1972) *J. Biol. Chem.* 247, 2205–2211.
- Jacobson, R. H., Zhang, X. J., DuBose, R. F. & Matthews, B. W. (1994) *Nature (London)* 369, 761–766.
- Kitamura, T., Onishi, M., Kinoshita, S., Shibuya, A., Miyajima, A. & Nolan, G. P. (1995) *Proc. Natl. Acad. Sci. USA* 92, 9146–9150.

Characterization of the Calmodulin-binding and of the Catalytic Domains of *Bordetella pertussis* Adenylate Cyclase*

XP-002100622

PD. 05-03-1989

(Received for publication, July 12, 1988)

Daniel Ladant†, Susan Michelson‡, Robert Sarfati§, Anne-Marie Gilles||, Roxane Predeleanu†, and Octavian Bărză†*

P. 4015-20-6

From the †Unité de Biochimie des Régulations Cellulaires, §Unité de Virologie Médicale, ¶Unité de Chimie Organique, and the ||Unité de Chimie des Protéines, Institut Pasteur, Paris, France

The structural organization of the low molecular mass form (43 kDa) of *Bordetella pertussis* adenylate cyclase was dissected taking advantage of the known sequence of the bacterial *cya* gene (Glaser, P., Ladant, D., Sezer, O., Pichot, F., Ullmann, A., and Danchin, A. (1988) *Mol. Microbiol.* 2, 19-30) and its low content of Trp and Met residues. Cleavage of the 43-kDa protein and of its complementary tryptic fragments (T25 and T18 peptides) with *N*-chlorosuccinimide and cyanogen bromide followed by sodium dodecyl sulfate-polyacrylamide gel analysis of digestion products allowed the following conclusions: (i) the catalytically active 43-kDa form of *B. pertussis* adenylate cyclase is within the first 400 residues of the protein encoded by the *cya* gene. T25 occupies the N-terminal domain of the protein (residues 1-235/237). Isolated T25 fragment exhibits a low but measurable enzymatic activity which indicates that it harbors the catalytic site; (ii) T18 which is the main calmodulin-binding domain, occupies the C-terminal segment of protein (residues 236/238-399) and is devoid of catalytic properties; (iii) the two complementary peptides T25 and T18 reassociated only in the presence of calmodulin, leading to significant recovery of the original activity. These results demonstrate that both fragments of the 43-kDa form of adenylate cyclase are essential for a high level of enzymatic activity.

Calmodulin-dependent adenylate cyclase is one of the factors implicated in the virulence of *Bordetella pertussis*, the causative agent of whooping cough (1). The existence of different molecular forms of this bacterial enzyme was reported by several groups (2-4). However, the size of the protein required for "invasiveness" of adenylate cyclase, i.e. its ability to enter eukaryotic cells, was difficult to ascertain in the absence of any structural data on the bacterial enzyme. Molecular cloning of *B. pertussis* adenylate cyclase gene showed that the protein is synthesized as a large precursor of 1706 residues. From genetic evidence it appeared that the calmodulin-stimulated catalytic activity resides in the 450 N-terminal amino acids of adenylate cyclase (5). On the other hand, biochemical studies showed that the low molecular mass form of *B. pertussis* adenylate cyclase isolated either from culture supernatants or from bacterial extracts consists of

three structurally related peptides of 50, 45, and 43 kDa (6). Limited proteolysis with trypsin of the adenylate cyclase/calmodulin complex first converted the 50- and 45-kDa peptides to the 43-kDa form, then the latter was cleaved into two complementary fragments of $M_r = 25,000$ and $M_r = 18,000$ called T25 and T18, which were still held together by calmodulin (CaM)¹ in an active, native-like structure (6). Cross-linking experiments with azido-CaM suggested that each fragment interacted with CaM when the adenylate cyclase-CaM complex was formed.

In this paper we further dissected the structural organization of the low molecular weight form of adenylate cyclase by cleavage at specific residues with chemical or enzymatic reagents. Our aim was to locate more precisely the CaM-binding and the catalytic sites of *B. pertussis* adenylate cyclase. We identified the T25 peptide as the N-terminal domain of adenylate cyclase which harbors the catalytic site, whereas the C-terminal T18 peptide is mainly involved in the binding of CaM.

MATERIALS AND METHODS

Chemicals—Adenine nucleotides were from Boehringer Mannheim. Bovine brain CaM, TPCK-trypsin, and soybean trypsin inhibitor were from Sigma. Methyl 4-azidobenzimidate was purchased from Pierce Chemical Co. Urea (fluorimetrically pure) was a product of Schwartz/Mann. [³²P]H₂PO₄ (carrier-free), [α -³²P]ATP (3000 Ci/mmol), [³H]cAMP (40 Ci/mmol) and Na¹²⁵I (1000 Ci/mmol) were obtained from the Radiochemical Centre, Amersham (United Kingdom). VU-8 CaM, made according to Craig *et al.* (7), was immobilized on CNBr-activated Sepharose 4B (a product of Pharmacia LKB Biotechnology Inc.) as described by Haiech *et al.* (8). Bovine brain CaM was azidated essentially according to Andreassen *et al.* (9) as described by Zurini *et al.* (10). Azido-CaM was iodinated by the chloramine-T method at room temperature to a specific activity of 0.02 mol of iodine/mol of azido-CaM. Bz₂ATP was synthesized as described by Williams and Coleman (11). [γ -³²P]Bz₂ATP (200 cpm/pmol) was prepared according to Glynn and Chappell (12). Synthetic peptide corresponding to residues 235-254 of *B. pertussis* adenylate cyclase (5) was obtained by a solid phase method on chloromethyl-substituted polystyrene resin cross-linked by 1% divinyl benzene. The residues were added using the symmetrical anhydride coupling method (13, 14). Cleavage from the resin and removal of all remaining protecting groups were accomplished by a low/high fluorhydric acid procedure (15).

Purification, Iodination, and Assay of Adenylate Cyclase—Adenylate cyclase was extracted from bacterial cells (phase I, type strain

* This work was supported by Grant URA 150 (UA 1129) from the Centre National de la Recherche Scientifique. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

** To whom correspondence should be addressed.

¹ The abbreviations used are: CaM, calmodulin; TPCK, L-1-tosylamido-2-phenylethyl chloromethyl ketone; NCS, *N*-chlorosuccinimide; CNBr, cyanogen bromide; Bz₂ATP, 3'-O-(4-benzoyl)benzoyladenine 5'-triphosphate; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; EGTA, (ethylenedis(oxyethyl)enitrilo)tetraacetic acid. P₂₃₅₋₂₅₄ is a synthetic peptide corresponding to residues 235-254 of *B. pertussis* adenylate cyclase. The sequence of P₂₃₅₋₂₅₄ is: Arg-Glu-Arg-Ile-Asp-Leu-Leu-Trp-Lys-Ile-Ala-Arg-Ala-Gly-Ala-Arg-Ser-Ala-Val-Gly.

ATCC 9797), with 4 M urea in buffer A (50 mM Tris-HCl, pH 8, 6 mM MgCl₂, and 0.1% Triton X-100) as described previously (16) and then purified in a single step chromatography on VU-8 CaM-Sepharose (8). Enzyme was iodinated by the chloramine-T method at room temperature to a specific activity of about 0.08–0.25 mol of iodine/mol of protein. Adenylate cyclase activity was measured in 50 mM Tris-HCl, pH 8, containing 6 mM MgCl₂, 0.1 mM CaCl₂, 0.1 μ M bovine brain CaM, 0.5 mg/ml bovine serum albumin, and 2 mM [α -³²P]ATP (5 \times 10⁶ cpm/assay) (16). One unit of adenylate cyclase activity corresponds to 1 μ mol of cAMP formed in 1 min at 30 °C and pH 8.

Preparation of a Homogeneous Low Molecular Mass Form (43 kDa) of B. pertussis Adenylate Cyclase—Iodinated adenylate cyclase (20 μ g/ml in buffer A, 5 \times 10⁵ cpm) was supplemented with 2.5 nmol of bovine brain CaM and 0.2 μ g/ml TPCK-trypsin. After 10 min of incubation at 4 °C (enzyme activity declined by less than 10%), 1 μ g/ml soybean trypsin inhibitor was added, and then the reaction mixture was loaded onto a 40 \times 1-cm column of Ultrogel AcA 44 equilibrated with buffer A. Elution was done at a flow rate of 4 ml/h. Fractions containing adenylate cyclase activity were pooled and precipitated with 5 volumes of cold acetone (–10 °C). Precipitated proteins were separated by centrifugation, solubilized with 9% formic acid, then loaded onto a column of 1 \times 66 cm of Sephadex G-100 equilibrated with 9% formic acid. Elution was done at a flow rate of 6 ml/h, and fractions of 1.1 ml were collected. After removal of formic acid by evaporation, samples were analyzed for purity by SDS-polyacrylamide gel and autoradiography.

Photoaffinity Labeling—Photoaffinity labeling experiments were performed at 4 °C for times varying between 1 and 30 min. The reaction mixture containing unlabeled adenylate cyclase and azido-¹²⁵I-CaM or [γ -³²P]Bz₂ATP was irradiated with a "long wave" mercury lamp (mineral light UVSL 58) positioned at 5 cm from the samples. Samples were then run on a 12.5 or 17.5% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) as described by Laemmli (17). Gels were either dried or fixed in 10% acetic acid and further washed with 25% isopropyl alcohol and 10% acetic acid. For detection of radiolabeled peptides on SDS-PAGE, dried gels were exposed at –80 °C to Kodak X-Omat AR films for 4–48 h with intensifying screens.

Reactivation of Adenylate Cyclase after SDS-PAGE—Iodinated adenylate cyclase or its tryptic fragments separated by SDS-PAGE were sliced from the gel with a razor blade, rehydrated in case of dried gels, and washed extensively with 25% isopropyl alcohol and then with 10% methanol to remove SDS. Slices in Eppendorf tubes were then soaked in 0.8 ml of 8 M urea in 50 mM Tris-HCl, pH 8, containing 0.1 mM CaCl₂ and 1% Nonidet P-40, for 18 h at 37 °C. Urea solutions containing different peptides were supplemented, or not, with bovine brain CaM and then dialyzed extensively against buffer A and assayed for adenylate cyclase activity as described above.

Cleavage of Adenylate Cyclase or of Its Tryptic Fragments with N-Chlorosuccinimide and Cyanogen Bromide—Peptide bands identified by SDS-PAGE and autoradiography were sliced from the gels, rehydrated when necessary, and then washed with either urea/water/acetic acid (1:1:1, w/v/v) (Trp cleavage) or 70% formic acid (Met cleavage) for 20 min. Peptides were cleaved by immersion of the gel slices at room temperature or at 37 °C in 15 mM NCS in urea/water/acetic acid for 30 min (Trp cleavage) or in 4% CNBr in 70% formic acid for 90 min (Met cleavage) (18–20). After several washings with water the slices were equilibrated in 10% glycerol, 5% β -mercaptoethanol, 2% SDS, and 62 mM Tris-HCl, pH 6.8, with several changes and loaded on a resolving gel. After SDS-PAGE the gels were dried and autoradiographed.

Analytical Procedures—Amino acid analysis was performed on a Biotronik amino acid analyzer LC 5001 using a single column procedure (21). Fluorescence measurements were performed on a Perkin-Elmer LS-5B luminescence spectrometer thermostated at 25 °C, using a 1 \times 1 UV-grade quartz cuvette (sample volume of 2 ml). Emission spectra of P_{235–254} peptide (λ_{exc} = 290 nm) in the presence or absence of Ca²⁺ or EGTA were recorded from 300 to 450 nm. Electrophoresis under nondenaturing conditions was performed according to Williams and Reisfeld (22) using a 13.75% resolution gel.

RESULTS

Amino Acid Composition of the 43-kDa Form of B. pertussis Adenylate Cyclase—Amino acid analysis of purified preparations of bacterial adenylate cyclase is complicated by the fact that the enzyme consists of three structurally related peptides

of 50, 45, and 43 kDa, as judged by SDS-PAGE. The proportion of these peptides varies greatly from one preparation to another. To circumvent this complicating factor we converted 50- and 45-kDa forms to the 43-kDa peptide by mild digestion with trypsin (100:1, w/w ratio) at 4 °C in the presence of a 5-fold excess of CaM over adenylate cyclase. The 43-kDa peptide was then separated from CaM and low molecular mass protein fragments by gel filtration under denaturing conditions. This procedure generated a single species of adenylate cyclase suitable for amino acid composition analysis. Alternatively, some preparations of adenylate cyclase were homogeneous enough (8) to be suitable for amino acid analysis. Table I shows the amino acid composition of the 43-kDa peptide as determined from two separate preparations of enzyme. The third column in Table I indicates the amino acid composition of the peptide encoded by the first 399 amino acids (M_r = 42,939) of adenylate cyclase as predicted from the nucleotide sequence of the *cya* gene. The number of 399 residues was arbitrarily chosen based on the closest M_r value to 43,000 and the fact that Arg-399 is a potential site of tryptic cleavage. With few exceptions (Met and Arg) the numbers of all other residues were remarkably similar. From these data one may assume that the low molecular weight form of *B. pertussis* adenylate cyclase may, indeed, correspond to the first 399 residues of the large 1706-residue precursor. However, without determination of the N-terminal sequence of the protein (our attempts to identify it failed, probably due to the fact that the N-terminal residue is blocked) other probes had to be taken into consideration.

Selective Tryptophanyl and Methionyl Peptide Bond Cleavage of 43-kDa, T25, and T18 Peptides—According to the primary structure of *B. pertussis* adenylate cyclase as deduced from the nucleotide sequence of *cya* gene corresponding to residues 1–399, the partial or complete cleavage of 43-kDa peptide at Trp-69 and Trp-242 should generate five fragments (Fig. 1B). This was indeed the case, as shown by the SDS-PAGE pattern of the NCS-digested iodinated protein (Fig. 2). The molecular weights of fragments calculated from a calibration curve of five molecular weight markers (Fig. 1B) corresponded within experimental error to the expected values, with a single exception. The smallest peptide fragment

TABLE I
Amino acid composition of *B. pertussis* adenylate cyclase (43-kDa protein)

Residue	Measured ^a	Deduced from DNA sequence (residues 1–399)
Cys	0	0
Asn + Asp	42.8	42
Thr	16.8	18
Ser	24.0	28
Gln + Glu	37.1	38
Pro	12	15
Gly	42.6	41
Ala	48.1 ^b	50
Val	32.4	32
Met	3.4	7
Ile	15.1	15
Leu	28.8	29
Tyr	9.9	12
Phe	13.9	14
His	9.2	9
Lys	13.1	15
Arg	38.1	32
Trp	ND ^c	2

^a The values given are uncorrected and determined after 20 h hydrolysis in 6 N HCl at 110 °C.

^b Arbitrarily taken as reference for M_r = 43,000.

^c ND, not determined.

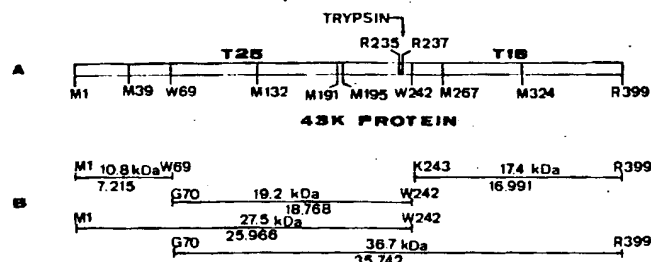


FIG. 1. Working model of the structural organization of the low molecular mass form of *B. pertussis* adenylate cyclase (43-kDa peptide). The positions of Met and Trp residues in the 43-kDa peptide, as deduced from the nucleotide sequence of the first quarter of *B. pertussis* *cya* gene, are indicated in A. The fragments resulting from partial or complete cleavage of adenylate cyclase after treatment with NCS, as well as the calculated M_r of fragments, are indicated in B. The molecular mass determined by SDS-PAGE using five proteins of known molecular weight is also indicated above each fragment.

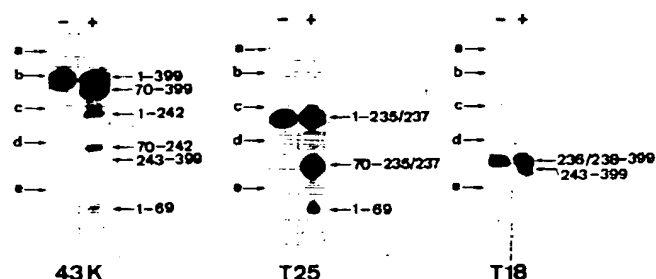


FIG. 2. Cleavage of 43-kDa, T25 and T18 peptides with *N*-chlorosuccinimide. 125 I-Adenylate cyclase (0.2 units, about 4×10^5 cpm) in 100 μ l of buffer A was supplemented with 0.4 μ M bovine brain CaM and 15 ng of TPCK-trypsin. After 10-min incubation at 4 $^{\circ}$ C, proteolysis was stopped with 0.1 μ g of soybean trypsin inhibitor and aliquots of 5–20 μ l were run on a 12.5% SDS-PAGE. Bands corresponding to uncleaved 43-kDa peptide, T25 and T18 were sliced from the gel and subjected (+) or not (–) to treatment with NCS as described under "Materials and Methods." Samples were run again on a 17.5% SDS-PAGE, and then the gels were dried and autoradiographed. The molecular weight markers (arrows on the left) were from top to bottom: bovine serum albumin, 67,000 (a); ovalbumin, 43,000 (b); carbonic anhydrase, 30,000 (c); soybean trypsin inhibitor, 20,300 (d); and lysozyme, 14,400 (e). The assignment of different peptides (arrows on the right) was made on the basis of their molecular weights and the primary structure (residues 1–399) of adenylate cyclase as deduced from the nucleotide sequence of the *cya* gene.

corresponding to residues 1–69 gave systematically higher molecular weight values than expected. This seems to be due to an anomalous migration of the 1–69 peptide since its complementary fragment (residues 70–399) had a molecular weight very close to the expected value.

Chemical cleavage at the tryptophanyl peptide bond of complementary T25 and T18 peptides gave a characteristic pattern which allowed assignment of T25 to the N-terminal domain of adenylate cyclase and T18 to the C-terminal domain of the enzyme. As shown in Fig. 2, T25 was cleaved by NCS into two fragments of apparent M_r of 16,600 and 10,800, respectively. The smallest of these migrated in SDS-PAGE exactly like the smallest peptide resulting from NCS cleavage of the 43-kDa peptide, which was previously assigned to residues 1–69. After NCS treatment, T18 gave a faint band of 16.5 kDa differing by only 1 kDa from the parent peptide. This indicates that the site of tryptic cleavage of the 43-kDa form of adenylate cyclase is slightly upstream of Trp-242. From the primary structure of the 43-kDa peptide, as deduced from the DNA sequence, it is reasonable to assign the site of

trypsin cleavage to Arg-235 or Arg-237 (Fig. 1A).

To further substantiate the location of T25 and T18 fragments in the primary structure of the *B. pertussis* adenylate cyclase 43-kDa peptide, these fragments were subjected to CNBr cleavage. The digestion pattern was more complex than in the case of NCS cleavage, due to the presence of seven methionine residues in the 43-kDa peptide. However, T18, which contains only 2 methionine residues, should have had a much simpler CNBr-cleavage pattern. This was indeed the case (Fig. 3).

Cross-linking of T25 and T18 Peptides to Azido- 125 I-CaM and Selective Cleavage with NCS—Unlabeled adenylate cyclase complexed to azido- 125 I-CaM was submitted to trypsin digestion and then photolysed. The cross-linked products of 63, 45, and 38 kDa corresponded to the 43-kDa, T25 and T18 peptides, respectively, covalently bound to azido- 125 I-CaM (6). Since CaM has no Trp residues, NCS cleavage of T25-CaM complex should tell us which segment of T25, situated before or after Trp-69, is involved in CaM binding. As shown in Fig. 4 (lane 5), after NCS cleavage of T25-CaM cross-linked product, a new radioactive band of 37 kDa appeared which seems to indicate that the segment of T25 responsible for CaM binding is situated after Trp-69. The cross-linked product of T18 and CaM did not generate any visible cleavage product after treatment with NCS.

Reconstitution of Catalytically Active Forms of Adenylate Cyclase from Fragments Resulting from Tryptic Cleavage of 43-kDa Peptide—Even after complete cleavage by trypsin of the CaM-complexed 43-kDa peptide, adenylate cyclase retained 70–80% of its original activity. Gel permeation chromatography of the CaM-T25-T18 complex revealed the same hydrodynamic properties as the undigested CaM-adenylate cyclase complex (6). Upon exposure to 8 M urea for 1 h followed by dilution, the CaM-T25-T18 complex was found to be inactive. However, when the trypsin-digested adenylate cyclase-CaM complex was exposed to 8 M urea and dialyzed against buffer A, 50% of the initial activity was recovered. Gel permeation chromatography of reactivated CaM and adenylate cyclase fragments again displayed the same hydrodynamic properties as the undigested or digested/undenatured CaM-adenylate cyclase complex (Fig. 5). These data strongly suggest that T25 and T18 refold after removal of urea and reassociate in the presence of CaM to yield a ternary T25-T18-CaM complex which is enzymatically active.

Since it has been shown that adenylate cyclase can be

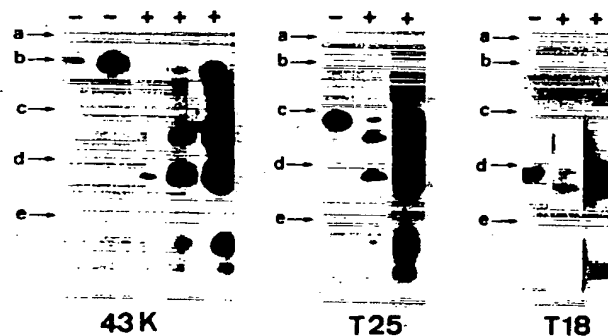


FIG. 3. Cleavage of 43-kDa, T25, and T18 peptides with cyanogen bromide. Iodinated adenylate cyclase was partially cleaved with TPCK-trypsin as described in the legend of Fig. 2. Peptide bands sliced from the gel were subjected (+) or not (–) to treatment with CNBr as described under "Materials and Methods." Samples were run again on a 17.5% SDS-PAGE, and then the gel was dried and autoradiographed using different exposure times to better visualize bands of different intensities. The molecular weight markers were the same as described in Fig. 2.

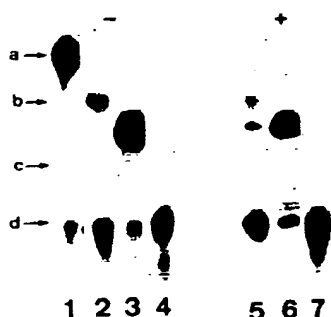


FIG. 4. Cross-linking of azido- ^{125}I -CaM to tryptic fragments of adenylate cyclase and cleavage of cross-linked products with *N*-chlorosuccinimide. Adenylate cyclase (0.2 units) was supplemented with 100 nM azido- ^{125}I -CaM and subjected to trypsin proteolysis as described in Fig. 2. Proteolysis was stopped with soybean trypsin inhibitor, then samples were irradiated for 1 min as described under "Materials and Methods," and run on a 10% SDS-PAGE. Radioactive bands corresponding to cross-linked products were sliced from the gel, subjected (+) or not (-) to treatment with NCS as described under "Materials and Methods," run again on a 12.5% SDS-PAGE, and then autoradiographed. Lane 1, 43-kDa peptide cross-linked to azido- ^{125}I -CaM; lanes 2 and 5, T25 cross-linked to azido- ^{125}I -CaM; lanes 3 and 6, T18 cross-linked to azido- ^{125}I -CaM; lanes 4 and 7, azido- ^{125}I -CaM. Arrows indicate the same molecular weight markers (a to d) as in Figs. 2 and 3.

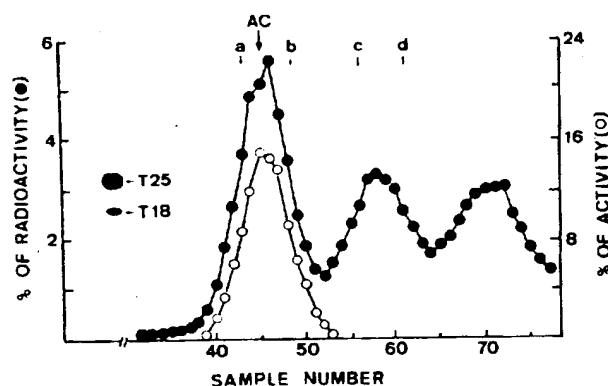


FIG. 5. Gel permeation chromatography of trypsin-cleaved adenylate cyclase-CaM complex after urea denaturation and renaturation. ^{125}I -adenylate cyclase (0.09 units; specific activity 3 units/ 10^7 cpm ^{125}I) in buffer A containing 0.6 μM CaM was digested for 10 min with 75 ng of trypsin at 4 $^\circ\text{C}$. The sample (residual specific activity, 2 units/ 10^7 cpm ^{125}I) was diluted in 200 μl of buffer A containing 8 M urea, 20 μg of soybean trypsin inhibitor, and 20 μg of CaM and incubated 1 h at 30 $^\circ\text{C}$. After an extensive dialysis against buffer A at 4 $^\circ\text{C}$, (recovered adenylate cyclase activity, 0.88 units/ 10^7 cpm ^{125}I), the mixture was loaded onto an Ultrogel AcA 44 column (0.6 \times 26 cm) equilibrated in buffer A containing 20 $\mu\text{g}/\text{ml}$ soybean trypsin inhibitor. Fractions of 0.1 ml were collected at a flow rate of 0.5 ml/h and analyzed for enzyme activity (O) and radioactivity (●). Inset shows an autoradiograph after a 12.5% SDS-PAGE of fractions corresponding to the peak of activity. The four molecular weight markers were: bovine serum albumin, 67,000 (a); ovalbumin, 43,000 (b); *Escherichia coli* adenylate kinase, 23,500 (c); and soybean trypsin inhibitor, 20,300 (d). AC designates the native adenylate cyclase-CaM complex.

reactivated in a high (about 50%) proportion after SDS-PAGE (6), we performed reconstitution experiments with T25 and T18 isolated by electrophoresis under denaturing conditions (see "Materials and Methods"). T18 renatured alone or in association with CaM did not show any cyclase activity. T25 exhibited a very low but measurable, enzyme activity (approximately 0.1–0.2% of the uncleaved protein at saturating CaM concentration), irrespective of the presence or absence

TABLE II

Adenylate cyclase activity of T25 and T18 peptides renatured in the presence or absence of CaM after SDS-PAGE

Reactivation of adenylate cyclase after SDS-PAGE is described under "Materials and Methods." Adenylate cyclase activity was determined in the presence of 100 nM CaM and 1 mM ATP.

Peptide	CaM (1 μM) during renaturation experiments	Specific activity	Recovery of activity ^a
		units/ 10^7 cpm	%
43 kDa	Present	2.83	54
T25	Present	0.012	0.22
T18	Present	0.000	0.00
T25 + T18	Absent	0.026	0.50
T25 + T18	Present	0.98	18.6

^a The specific activity of ^{125}I -adenylate cyclase before trypsin digestion was taken as 100% (5.25 units/ 10^7 cpm).

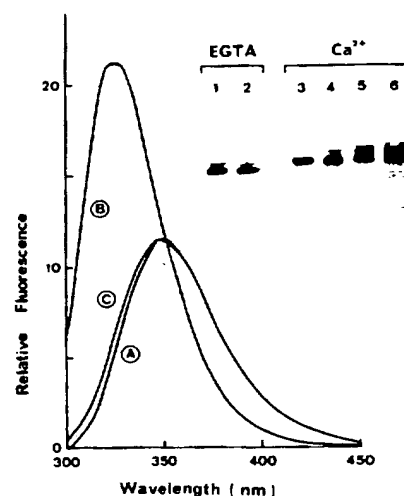


FIG. 6. Ca^{2+} -dependent binding of synthetic $\text{P}_{235-254}$ peptide to CaM as shown by fluorescence and gel retardation analysis (inset). Emission spectra of $\text{P}_{235-254}$ (5 μM) in 20 mM Tris-HCl, pH 7.5, 0.1 M NaCl, and 0.2 mM CaCl_2 were recorded from 300 to 450 nm. A, synthetic peptide alone; B, synthetic peptide plus 5 μM bovine brain CaM; C, same as B, after addition of 2 mM EGTA. Inset shows the electrophoretic behavior of CaM (5 μg) under native conditions in the presence of $\text{P}_{235-254}$ peptide and excess Ca^{2+} (0.2 mM) or EGTA (2 mM). Lanes 1 and 3, no peptide present; lane 4, peptide/CaM (molar ratio), 0.5:1; lane 5, peptide/CaM, 1:1; lanes 2 and 6, peptide/CaM, 4:1.

of CaM during renaturation. Assuming that CaM activation of *B. pertussis* adenylate cyclase is between 20- and 50-fold, the catalytic activity of T25 might be estimated as between 2 and 10% of that of the whole 43-kDa peptide. When a urea solution containing complementary fragments was mixed with CaM prior to the dialysis step, adenylate cyclase activity was recovered to a significant level (Table II). If CaM was omitted during renaturation of T25 and T18, recovery of catalytic activity was considerably lower.

From these results one may assume that the catalytic site of adenylate cyclase is located on the T25 domain, but full catalytic activity is expressed by a structure in which T25 and T18 are "bridged" together by CaM. The values of the apparent K_m for ATP of the native 43-kDa form, of the enzyme renatured after SDS-PAGE or of the reconstituted T25-T18-CaM complex were between 0.6 and 0.8 mM.

Interaction of the Synthetic Peptide $\text{P}_{235-254}$ with Bovine Brain Calmodulin—Although it is difficult to identify a typical CaM-binding site on the sequence of *B. pertussis* adenylate cyclase like the site described for skeletal or smooth muscle

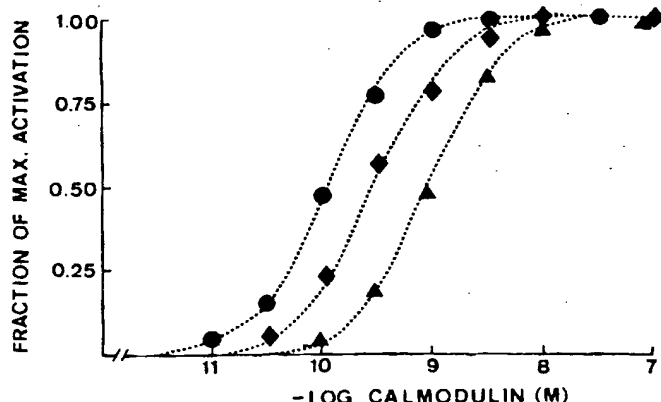


FIG. 7. Reversal of the CaM-dependent activation of *B. pertussis* adenylate cyclase by synthetic peptide P₂₃₅₋₂₅₄. CaM diluted to desired concentrations was added in 5- μ l aliquots to the reaction mixture described under "Materials and Methods" (●). When present, P₂₃₅₋₂₅₄ was at 0.5 μ M (◆) and 5 μ M (▲). The reaction was initiated by the addition of adenylate cyclase (0.5 ng/assay). The fraction of maximal activation was calculated as $(V - V_0)/(V_{max} - V_0)$, where V_0 and V_{max} are the reaction rates in the absence of CaM or in the presence of saturating concentrations of CaM; V is the reaction rate at a given concentration of CaM. The dissociation constant of synthetic peptide-CaM complex (580 nm) was estimated from the midpoints of the activation curves as described by Erickson-Viitanen and De Grado (26).

myosin light chain kinase (23, 24), the sequence corresponding to residues 235–254 seems to be a candidate for such a CaM-binding site because: (i) this sequence belongs to the N terminus of the T18 peptide and joins the T25 fragment of adenylate cyclase via Arg-235 (or Arg-237); (ii) azido-¹²⁵I-CaM preferentially photolabels the T18 tryptic domain of bacterial enzyme; (iii) this sequence has a Trp in its first half and shows a strong predominance of basic and hydrophobic residues (25).

The fluorescence emission spectrum of synthetic peptide P₂₃₅₋₂₅₄ had a maximum at 350 nm (Fig. 6) and was not influenced by the presence of Ca²⁺ or EGTA. Addition of CaM in a 1:1 ratio to P₂₃₅₋₂₅₄ in the presence of 0.2 mM Ca²⁺ caused a 1.82-fold increase in total fluorescence intensity and a shift of the spectral maximum to 325 nm. EGTA promoted dissociation of the CaM-P₂₃₅₋₂₅₄ complex. The Ca²⁺-dependent CaM-binding properties of P₂₃₅₋₂₅₄ were confirmed by electrophoresis of CaM-peptide complex under nondenaturing conditions. In the presence of Ca²⁺, P₂₃₅₋₂₅₄ retarded the migration of CaM relative to that achieved in the presence of EGTA (Fig. 6, inset). The synthetic peptide P₂₃₅₋₂₅₄ reversed CaM-dependent activation of *B. pertussis* adenylate cyclase (Fig. 7). The relatively high K_d value of P₂₃₅₋₂₅₄ for CaM (580 nM) as compared to other CaM-binding sequences (23–29) may indicate that it represents only part of a high affinity binding domain.

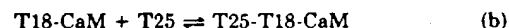
DISCUSSION

Calmodulin regulates the activity of a large number of enzymes with various catalytic functions and biological significance. The molecular basis of the interaction of CaM with target proteins is beginning to be understood on the basis of two independent approaches: (i) isolation of CaM-binding fragments from enzymatic or chemical digests of target enzymes (25, 28, 30); (ii) synthesis of various peptides able to interact with CaM in a Ca²⁺-dependent fashion (27, 31, 32). A consensus emerging from these studies is that the activator-binding domains of CaM-dependent enzymes show only a

limited degree of strict sequence homology but form basic amphiphilic helices, in agreement with earlier observations (33–35).

An important advantage in approaching the location of catalytic and CaM-binding domains of *B. pertussis* adenylate cyclase is the high specific activity of the enzyme, the ease of its reactivation after different treatments including SDS-PAGE, and knowledge of the nucleotide sequence of the *B. pertussis* *cya* gene. Thus, we were able to show in agreement with genetic arguments (5) that the catalytically active, low molecular weight form of adenylate cyclase (present in bacterial culture supernatant) corresponds to the first quarter of the coding sequence of *B. pertussis* *cya* gene. Selective cleavage of the protein by trypsin yielded two fragments, one corresponding to the N-terminal domain carrying the catalytic site (T25), the second corresponding to the C-terminal domain, which is responsible for the interaction with CaM (T18).

Location of the CaM-binding site of *B. pertussis* adenylate cyclase around Trp-242 was suggested by photoaffinity labeling experiments of protein with azido-¹²⁵I-CaM, as well as by the fact that synthetic peptide corresponding to residues 235–254 interacted with CaM in a Ca²⁺-dependent manner. However, the affinity of the synthetic peptide for CaM, compared to other well known CaM-binding sequences, was unexpectedly low to make certain such an assignment. Site-directed mutagenesis of Trp-242 with Arg, Gly, and Asp was helpful in confirming the location of CaM-binding site of adenylate cyclase. All three mutants were fully active, but affinity for CaM was decreased by a factor varying between 30 and 1000.² One might envisage that the spatial relationship of the catalytic site and the CaM-binding site of *B. pertussis* adenylate cyclase differs from that of other CaM-dependent enzymes. The particular organization of the CaM-binding site of *B. pertussis* adenylate cyclase might also explain why the reconstitution of catalytically active species of adenylate cyclase from isolated T25 and T18 fragments requires the presence of CaM. If we admit the following sequence of interactions:



where the binary complex is of low stability and the ternary complex is of high stability, the limiting step in the formation of the catalytically active species is the formation of the T18-CaM complex.

Attempts to locate the catalytic site on T25 peptide using the photoactivatable ATP analog B₂ATP failed. Although irreversible binding of [γ -³²P]B₂ATP to the enzyme upon photolysis was demonstrated, the photolabel was too unstable under conditions of protein cleavage. By comparing the sequence situated between residues 54 and 70 of *B. pertussis* adenylate cyclase with the well known motif Gly-X-X-X-X-Gly-Lys-Thr-(Ser) in ATP-binding proteins (36–38), one may surmise that the polyphosphate binding site in *B. pertussis* adenylate cyclase is located near Lys-65. This lysine belongs to a β -turn segment, as predicted from the primary structure of the protein using the algorithm of Garnier *et al.* (39). Similar sequences in ATP-binding proteins form glycine-rich flexible loops terminated by a lysine which, by changing its conformation and location, may control access to the MgATP site (36). Replacement of Lys-65 in *B. pertussis* adenylate cyclase with Gln yielded mutant enzyme exhibiting only 0.1% of catalytic activity of wild type protein and intact CaM-binding properties.²

² Glaser, P., Elmaoglou-Lazaridou, A., Krim, E., Ladant, D., Bärzu, O., and Danchin, A. (1989) *EMBO J.*, in press.

Acknowledgments—We thank A. Ullmann for helpful discussion and constructive criticisms, A. Danchin, P. Glaser, and M. Mock for fruitful comments during preparation of the manuscript as well as for communicating data prior to publication, O. Siffert for synthesis of P₂₃₅₋₂₅₄ peptide, J. Bellalou for providing bacterial cultures, T. Rose for analysis of the secondary structure of adenylate cyclase, A. Chafotte for useful advices in fluorescence experiments, J. Haiech (Centre de Recherches de Biochimie Macromoléculaire, Montpellier) for the generous gift of VU-8 CaM, and M. Ferrand for expert secretarial help.

REFERENCES

- Weiss, A. A., and Hewlett, E. L. (1986) *Annu. Rev. Microbiol.* **40**, 661-686
- Masure, H. R., Shattuck, R. L., and Storm, D. R. (1987) *Microbiol. Rev.* **51**, 60-65
- Kessin, R. H., and Franke, J. (1986) *J. Bacteriol.* **166**, 290-296
- Hanski, E., and Farfel, Z. (1985) *J. Biol. Chem.* **260**, 5526-5532
- Glaser, P., Ladant, D., Sezer, D., Pichot, F., Ullmann, A., and Danchin, A. (1988) *Mol. Microbiol.* **2**, 19-30
- Ladant, D. (1988) *J. Biol. Chem.* **263**, 2612-2618
- Craig, T. A., Watterson, D. M., Prendergast, F. G., Haiech, J., and Roberts, D. M. (1987) *J. Biol. Chem.* **262**, 3278-3284
- Haiech, J., Predeleanu, R., Watterson, D. M., Ladant, D., Bellalou, J., Ullmann, A., and Bärzu, O. (1988) *J. Biol. Chem.* **263**, 4259-4262
- Andreasen, T. J., Keller, C. H., LaPorte, D. C., Edelman, A. M., and Storm, D. R. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78**, 2782-2785
- Zurini, M., Krebs, J., Penniston, J. T., and Carafoli, E. (1984) *J. Biol. Chem.* **259**, 618-627
- Williams, N., and Coleman, P. S. (1982) *J. Biol. Chem.* **257**, 2834-2841
- Glynn, I. M., and Chappell, J. B. (1964) *Biochem. J.* **90**, 147-149
- Hancock, W. S., Prescott, D. J., Vagelos, P. R., and Marshall, G. R. (1973) *J. Org. Chem.* **38**, 774-781
- Atherton, E., Caviezel, M., Fox, J., Harkiss, D., Over, H., and Scheppard, R. C. (1983) *J. Chem. Soc. I*, 65-73
- Tam, J. P., and Merrifield, R. B. (1985) *Int. J. Peptide Protein Res.* **26**, 262-273
- Ladant, D., Brezin, C., Alonso, J.-M., Crenon, I., and Guiso, N. (1986) *J. Biol. Chem.* **261**, 16264-16269
- Laemmli, U. K. (1970) *Nature* **227**, 680-685
- Shechter, Y., Patchornik, A., and Burnstein, Y. (1976) *Biochemistry* **15**, 5071-5075
- Nikodem, V., and Fresco, J. R. (1979) *Anal. Biochem.* **97**, 382-386
- Lischwe, M. A., and Ochs, D. (1982) *Anal. Biochem.* **127**, 453-457
- Hummel, B. C. W. (1959) *Can. J. Biochem. Physiol.* **37**, 1393-1399
- Williams, D. E., and Reisfeld, R. A. (1964) *Ann. N. Y. Acad. Sci.* **121**, 373-381
- Kennelly, P. J., Edelman, A. M., Blumenthal, D. K., and Krebs, E. G. (1987) *J. Biol. Chem.* **262**, 11958-11963
- Lukas, T. J., Burgess, W. H., Prendergast, F. G., Lau, W., and Watterson, D. M. (1986) *Biochemistry* **25**, 1458-1464
- James, P., Maeda, M., Fischer, R., Verma, A. K., Krebs, J., Penniston, J. T., and Carafoli, E. (1988) *J. Biol. Chem.* **263**, 2905-2910
- Erickson-Viitanen, S., and De Grado, W. F. (1987) *Methods Enzymol.* **139**, 455-478
- O'Neil, K. T., Wolfe, H. R., Jr., Erickson-Viitanen, S., and De Grado, W. F. (1987) *Science* **236**, 1454-1456
- Buschmeier, B., Meyer, H. E., and Mayr, G. W. (1987) *J. Biol. Chem.* **262**, 9454-9462
- Hanley, R. M., Means, A. R., Kemp, B. E., and Shenolikar, S. (1988) *Biochem. Biophys. Res. Commun.* **152**, 122-128
- Blumenthal, D. K., Takio, K., Edelman, A. M., Charbonneau, H., Titani, K., Walsh, K. A., and Krebs, E. G. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82**, 3187-3191
- De Grado, W. F., Erickson-Viitanen, S., Wolfe, H. R., Jr., and O'Neil, K. T. (1987) *Proteins* **2**, 20-33
- De Grado, W. F., Prendergast, F. G., Wolfe, H. R., Jr., and Cox, J. A. (1985) *J. Cell Biochem.* **29**, 83-93
- Weiss, B., Prozialeck, W., Cimino, M., Barnette, M. S., and Wallace, T. L. (1980) *Ann. N. Y. Acad. Sci.* **356**, 319-345
- Barnette, M. S., Daly, R., and Weiss, B. (1983) *Biochem. Pharmacol.* **32**, 2929-2933
- Malencik, D. A., and Anderson, S. R. (1983) *Biochem. Biophys. Res. Commun.* **114**, 50-56
- Moller, W., and Amons, R. (1985) *FEBS Lett.* **186**, 1-7
- Fry, D. C., Kubly, S. A., and Mildvan, A. S. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83**, 907-911
- Sternlicht, H., Yaffe, M. B., and Farr, G. W. (1987) *FEBS Lett.* **214**, 226-235
- Garnier, J., Osguthorpe, D. J., and Robson, B. (1978) *J. Mol. Biol.* **120**, 97-120



XP 000082630

6084 The EMBO Journal

8 (1989) March, no. 3, Eynsham, Oxford, GB

The EMBO Journal vol. 8 no. 3 pp. 967-972, 1989

C12N15/31

Identification of residues essential for catalysis and binding of calmodulin in *Bordetella pertussis* adenylate cyclase by site-directed mutagenesis

Philippe Glaser, Arétie Elmaoglou-Lazaridou¹,
Evelyne Krin, Daniel Ladant¹, Octavian Bârză¹
and Antoine Danchin

Unité de Régulation de l'Expression Génétique and ¹Unité de
Biochimie des Régulations Cellulaires, Institut Pasteur, 28 rue du
Docteur Roux, 75724 Paris Cedex 15, France

Communicated by A. Danchin

In order to identify molecular features of the calmodulin (CaM) activated adenylate cyclase of *Bordetella pertussis*, a truncated *cya* gene was fused after the 459th codon in frame with the α -*lacZ'* gene fragment and expressed in *Escherichia coli*. The recombinant, 604 residue long protein was purified to homogeneity by ion-exchange and affinity chromatography. The kinetic parameters of the recombinant protein are very similar to that of adenylate cyclase purified from *B. pertussis* culture supernatants, i.e. a specific activity $> 2000 \mu\text{mol/min mg}$ of protein at 30°C and pH 8, a K_m^{ATP} of 0.6 mM and a K_d for its activator, CaM, of 0.2 nM. Proteolysis with trypsin in the presence of CaM converted the recombinant protein to a 43 kD protein with no loss of activity; the latter corresponds to the secreted form of *B. pertussis* adenylate cyclase. Site-directed mutagenesis of residue Trp-242 in the recombinant protein yielded mutants expressing full catalytic activity but having altered affinity for CaM. Thus, substitution of an aspartic acid residue for Trp-242 reduced the affinity of adenylate cyclase for CaM > 1000 -fold. Substitution of a Gln residue for Lys-58 or Lys-65 yielded mutants with a drastically reduced catalytic activity ($\sim 0.1\%$ of that of wild-type protein) but with little alteration of CaM-binding. These results substantiated, at the molecular level, our previous genetic and biochemical studies according to which the N-terminal tryptic fragment of secreted *B. pertussis* adenylate cyclase (residues 1-235/237) harbours the catalytic site, whereas the C-terminal tryptic fragment (residues 235/237-399) corresponds to the main CaM-binding domain of the enzyme.

Key words: ATP-binding site/*Bacillus anthracis*/calmodulin-binding site/cyclolysin

Introduction

Bordetella pertussis adenylate cyclase is an intensely studied enzyme because of its potential role in the pathogenesis of whooping cough (Weiss and Hewlett, 1986). Two properties of adenylate cyclase from this organism make it an attractive model for structure-function relationship studies: (i) its activation by calmodulin (CaM) which is not known to occur in bacteria (Wolff *et al.*, 1980; Greenlee *et al.*, 1982); and (ii) its ability to enter eukaryotic cells causing unregulated synthesis of adenosine 3'-5'-monophosphate

(cAMP) and impairment of normal cellular function (Confer and Eaton, 1982; Hanski and Farfel, 1985). Delineation of the mechanism of activation and catalysis in *B. pertussis* adenylate cyclase would provide the first instance where the active site of this heterogeneous group of enzymes could be characterized.

Using the interaction between adenylate cyclase and CaM as a tool we have recently cloned the corresponding gene from *B. pertussis*, determined its nucleotide sequence and expressed it in *Escherichia coli* (Glaser *et al.*, 1988a). We have found that the protein is synthesized as a large bifunctional precursor form of 1706 amino acid residues, endowed with adenylate cyclase and haemolytic activity. The protein has been named cyclolysin to indicate this fact. Its secretion mechanism involves the haemolytic carboxy-terminal end (Glaser *et al.*, 1988b) in conjunction with the product of three genes located downstream from the adenylate cyclase gene. In *B. pertussis*, the large precursor form is processed to low molecular mass forms of 43, 45 and 50 kD present in *B. pertussis* culture supernatant (Shattuck *et al.*, 1985; Ladant *et al.*, 1986). The secreted adenylate cyclase(s) corresponds to the amino-terminal domain of the large precursor form. By combining genetic and biochemical information (Glaser *et al.*, 1988a; Ladant, 1988; Ladant *et al.*, 1989), we showed that the N-terminal tryptic fragment of the 43 kD form of adenylate cyclase (T25, residues 1-235/237) harbours the active site, whereas the C-terminal tryptic fragment (T18, residues 235/237-399) corresponds to the main CaM-binding domain of the enzyme. Furthermore, a synthetic peptide corresponding to residues 235-254 of adenylate cyclase was shown to bind to CaM in a Ca^{2+} -dependent manner (Ladant *et al.*, 1989).

In the present study, we examined the catalytic and CaM-binding properties of *B. pertussis* adenylate cyclase expressed in *E. coli* by a plasmid carrying a truncated *cya* gene which comprises the entire cyclase domain of cyclolysin fused to the α domain of β -galactosidase. Site-directed mutagenesis of Lys-58, Lys-65 and Trp-242 revealed the role of these residues in catalysis or in binding of CaM.

Results

Expression and purification of a truncated *B. pertussis* *cya* gene product in *E. coli*

Several clones deleted at the 3' end of the *cya4* gene were obtained as side products of the sequence analysis of the *B. pertussis* adenylate cyclase gene using the cyclone strategy. Clones encoding enzyme with a M_r close to that of adenylate cyclase purified from *B. pertussis* culture supernatants were transferred into the expression vector pTZ19 (Pharmacia, Uppsala, Sweden). One such clone, plasmid pDIA5202, gave the highest level of adenylate cyclase activity. As indicated in Figure 1, it fused sequences corresponding to the whole adenylate cyclase domain of

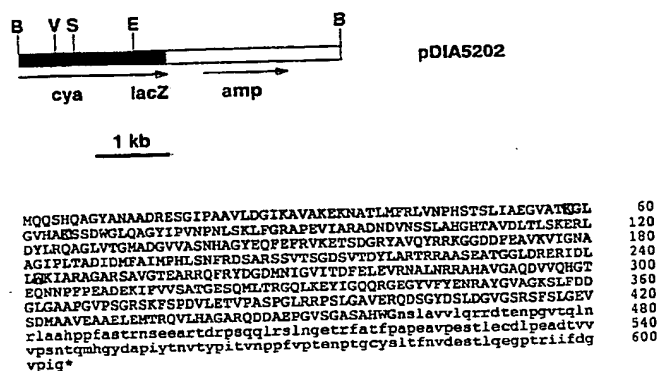


Fig. 1. Restriction map of plasmid pDIA5202 and deduced amino acid sequence of the recombinant adenylate cyclase. The 5'-terminal end of the cyclolysin gene has been fused in phase with a truncated *lacZ* gene containing a sequence specifying the α domain of β -galactosidase. This resulted in a chimeric protein containing the adenylate cyclase domain of cyclolysin (residues 1–399), the start of the haemolysin domain (residues 400–459) and the α domain of β -galactosidase (residues 460–604, in lower case letters). The Arg pair (399/400) is particularly sensitive to trypsin action resulting in formation of a 43 kd protein, identical to the small adenylate cyclase form present in *B. pertussis* supernatants. The black box corresponds to the inserted *B. pertussis* chromosomal DNA. The truncated *lacZ* gene is shaded. Restriction sites are represented as follows: B, *Bam*HI; E, *Eco*RI; V, *Eco*RV; S, *Sal*I. Amino acid residues are represented by the standard one-letter code. Residues submitted to site-directed mutagenesis (K58, K65 and W242) are boxed.

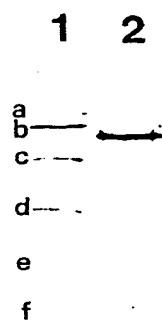


Fig. 2. SDS-PAGE (12.5%) of purified recombinant adenylate cyclase expressed in *E. coli*. Lane 1: M, markers: a, phosphorylase a, (94 000); b, bovine serum albumin (67 000); c, ovalbumin (43 000); d, carbonic anhydrase (30 000); e, soybean trypsin inhibitor (20 300); f, lysozyme (14 400). Lane 2: purified adenylate cyclase (2 µg of proteins). Gels were stained with Coomassie Blue.

cyclolysin and the α domain of β -galactosidase. Since the transcription and translation signals were identical in different plasmids, this higher activity was probably due to a better stability of the protein. This clone was therefore kept for further studies.

The maximum level of adenylate cyclase activity was reached when cells entered the early stationary phase, after which a decrease of activity was noted, probably due to the degradation of the protein. During the exponential phase of growth, the ratio of β -galactosidase to adenylate cyclase activity was constant, as expected for α -complementation (interaction between peptide α carried by the recombinant adenylate cyclase and the partially deleted β -galactosidase coded by the *lacZ*M15 gene). The two enzyme activities

Table I. Purification of recombinant *B. pertussis* adenylate cyclase expressed in *E. coli*

Step	Protein (mg)	Total activity (µmol/min)	Specific activity (µmol/min mg of protein)	Yield (%)
Whole cell extract	889	2410	2.7	100
8000 g sediment in 8 M urea	138	1968	14.2	82
DEAE-Sephacel chromatography	17.8	1032	60.0	43
CaM-agarose chromatography	0.35	738	2104	31

Bacteria from 2 l of culture were suspended in 50 ml of 20 mM K-phosphate (pH 7.4) then disrupted by ultrasound. The extract was centrifuged at 8000 g for 30 min and the sediment resuspended in 20 ml of 8 M urea in Tris-Triton buffer [50 mM Tris-HCl, pH 8 plus 0.1% (w/w) Triton X-100]. After removal of insoluble material by centrifugation at 13 500 g for 30 min, the 'urea extract' was loaded onto a DEAE-Sephacel column (20 mg of protein/ml of swollen gel) at a flow rate of 30 ml/h. The column was washed with 150 ml of Tris-Triton buffer for 1.5 h, then adenylate cyclase was eluted with 0.5 M NaCl in the same buffer. Fractions containing adenylate cyclase were supplemented with 5 mM CaCl_2 then purified on CaM-agarose as described by Monneron *et al.* (1988).

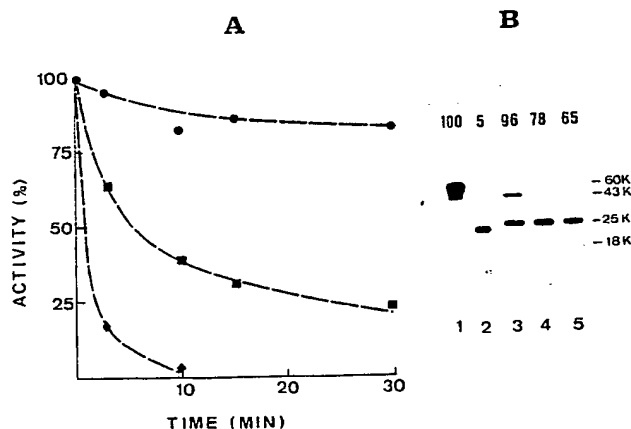


Fig. 3. Trypsin digestion of recombinant adenylate cyclase in the presence and absence of CaM. (A) Seven units of purified recombinant adenylate cyclase in 100 µl of Tris-Triton buffer (50 mM Tris-HCl, pH 8 plus 0.1% Triton X-100) supplemented with 1 mM CaCl_2 and 50 ng trypsin-trypsin were incubated at 4°C in the presence of 2 µM CaM (●, ■) or 0.5 mg/ml human serum albumin (♦). At different time intervals 5 µl aliquots were withdrawn and diluted in 200 µl of Tris-Triton buffer containing 2 µg of soybean trypsin inhibitor (●, ♦) or in 35 µl of 8 M urea in the same buffer (■). After 10 min at 4°C, samples in urea were diluted to 200 µl with Tris-Triton buffer, then residual adenylate cyclase activity (expressed as percentage of initial activity) was determined. Controls run in the absence of trypsin and diluted with Tris-Triton buffer or 8 M urea Tris-Triton showed no decrease in enzyme activity. (B) 0.1 units of pure [^{125}I]adenylate cyclase (7×10^5 c.p.m.) in 50 µl of Tris-Triton buffer containing 1 mM CaCl_2 were submitted to TPCK-trypsin (25 ng) proteolysis for 15 s (lane 3), 3 min (lane 4), 5 min (lane 2) or 7 min (lane 5), in the absence (lane 2) or in the presence of 0.15 µM CaM (lanes 3, 4 and 5). Proteolysis was stopped with an excess of soybean trypsin inhibitor (2 µg/sample). Residual adenylate cyclase activity (expressed as percentage of initial activity on the top of the gel lanes) was determined for each sample. The remaining samples were run onto 12.5% (w/v) SDS-PAGE and autoradiographed. The arrows to the right indicate the positions of the undigested protein (lane 1, 100% activity), the 43 kd form of adenylate cyclase, and the two complementary peptides resulted from trypsin digestion of *B. pertussis* adenylate cyclase (25 and 18 kd, respectively).

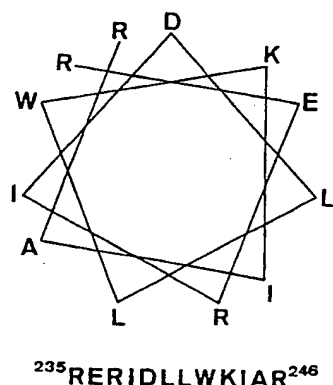


Fig. 4. A helix wheel representation of the putative CaM-binding sequence situated between residues 235 and 246 in *B. pertussis* adenylate cyclase.

Table II. Adenylate cyclase activity of crude extracts

Strain	Activity (nmol/min mg of protein)	
	10 nM CaM	2 μ M CaM
WT (S202)	2610	2540
W242D	87	1330
W242G	480	2220
W242R	1465	2430
W242V	2530	2820
K65Q	1.51	1.40
K58Q	0.55	0.59

Bacteria was disrupted by ultrasound and the extract was diluted with 8 M urea in Tris-Triton buffer to 'unmask' cryptic adenylate cyclase activity. K65Q and K58Q mutant extracts were analysed for enzyme activity using incubation times of 1 h instead of 5 or 10 min.

were recovered in the pellet fraction after cell breakage and centrifugation.

Formation of adenylate cyclase- β -galactosidase aggregates was found to be helpful in enzyme purification. Insoluble proteins once treated with urea in buffer solution contained >80% of the original adenylate cyclase activity. Chromatography on DEAE-Sephacel allowed removal of the chaotropic agent and recovery of a fully active enzyme by elution with salt. Finally, affinity chromatography on CaM-agarose yielded pure adenylate cyclase (Table I and Figure 2). The mol. wt of the protein calculated from a calibration curve is at 60 kd, somewhat lower than that expected from the deduced sequence of the protein (M_r = 64 995). Several faint bands, visible between 60 and 43 kd correspond most probably to products of cleavage of the recombinant protein by endogenous proteases. Antibodies directed against adenylate cyclase of *B. pertussis* (Monneron *et al.*, 1988) recognized these polypeptides in Western blots (data not shown).

Characterization of the *B. pertussis* adenylate cyclase expressed in *E. coli*

E. coli carrying the truncated gene of *B. pertussis* adenylate cyclase expressed not only an insoluble, but also a cryptic, form of the enzyme after growth at 37°C, most probably due to the fact that adenosine 5'-triphosphate (ATP) and/or

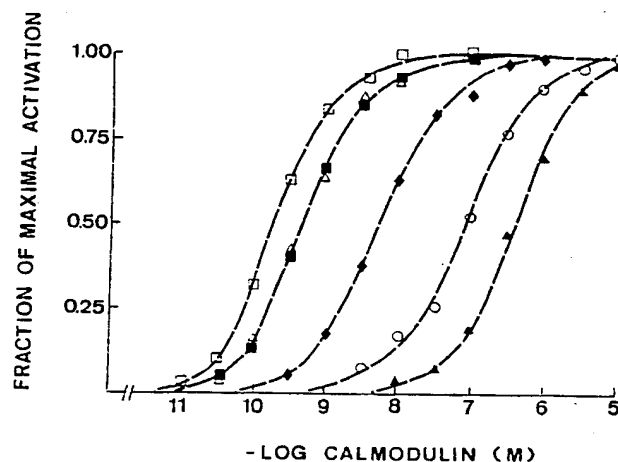


Fig. 5. Activation by calmodulin of wild-type (\square) and different mutant forms (W242V, \blacksquare ; W242R, \blacklozenge ; W242D, \blacktriangle ; W242G, \circ ; and K65Q, \triangle) of recombinant adenylate cyclase expressed in *E. coli*. Bovine brain calmodulin (100 μ M stock solution) was diluted to the desired concentrations with Tris-Triton buffer containing 1 mg/ml human serum albumin and then added in 5 μ l-aliquots to the reaction medium described in Materials and methods. CaCl_2 was present at 0.1 mM. The reaction was initiated by the addition of adenylate cyclase followed by 10-60 min incubation at 30°C. The fraction of maximal activation was calculated as $(V - V_0)/(V_{\max} - V_0)$ where V_0 and V_{\max} are the reaction rates in the absence of CaM or in the presence of saturating concentrations of CaM, respectively; V is the reaction rate at a given concentration of CaM. The wild-type and K65Q mutant enzymes were used as purified preparations; the Trp mutants were derived from crude bacterial extracts.

CaM were less accessible to the cyclase moiety of the complemented protein aggregates. Dilution of crude extracts with 8 M urea instead of buffer prior to activity assays unmasked this cryptic activity. Trypsin rapidly inactivated the recombinant adenylate cyclase in the absence of CaM as was the case with enzyme secreted from *B. pertussis* (Figure 3A). Proteolysis performed in the presence of CaM converted the 60 kd protein to a 43 kd form with no loss of activity. Further proteolysis of the 43 kd form yielded the complementary fragments T25 and T18 (Ladant, 1988), still maintained in an active native-like structure by CaM (Figure 3B). It is therefore clear that the α peptide of β -galactosidase, as well as residues between 400 and 458 of recombinant adenylate cyclase, could easily be removed from the protein by tryptic digestion with no significant loss of activity. The kinetic parameters of the recombinant protein are very similar to those of adenylate cyclase purified from *B. pertussis* culture supernatants, i.e. a specific activity >2000 μ mol/min mg of protein at 30°C and pH 8, a K_m for ATP of \sim 0.6 mM and a K_d for CaM of 0.2 nM in the presence of Ca^{2+} ions. Ethylene bis(oxyethylenetriole) tetraacetic acid (EGTA) decreased the affinity of recombinant adenylate cyclase for CaM by about two orders of magnitude, as was the case with enzyme secreted from *B. pertussis* (Ladant, 1988), while the maximum catalytic activity was increased by a factor of 1.7 (data not shown).

Site-directed mutagenesis of Trp-242 in recombinant adenylate cyclase

Attempts to identify the CaM-binding site of *B. pertussis* adenylate cyclase by sequence comparison with well-characterized CaM-activated enzymes was rather inconclusive.

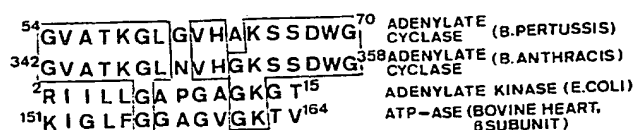


Fig. 6. Alignment of the putative polyphosphate-binding sites of *B. pertussis* adenylate cyclase with sequences belonging to well-characterized ATP-binding enzymes.

However, the sequence situated around Trp-242 in the bacterial enzyme (Figure 4) had some features characteristic of CaM-binding peptides such as presence of basic and hydrophobic residues forming amphiphilic α -helix structures (Blumenthal *et al.*, 1985; Buschmeier *et al.*, 1987; Kennelly *et al.*, 1987; O'Neil *et al.*, 1987; Glaser *et al.*, 1988a; Hanley *et al.*, 1988; James *et al.*, 1988). Since Trp is frequently present in CaM-binding sequences, Trp-242 was chosen as the first target for analysis of *B. pertussis* adenylate cyclase by site-directed mutagenesis.

As shown in Table II replacement of Trp-242 in wild-type adenylate cyclase with Val, Arg, Gly or Asp (W242V, W242R, W242G and W242D mutants) had little effect on enzyme activity in crude extracts, as long as CaM concentration was well above saturation. However, at 10 nM CaM (concentration at which the activator is almost saturating for the wild-type enzyme) the four mutants displayed considerable differences in activity: mutant W242V was the least affected (90% of maximal activity at 10 nM CaM), whereas mutant W242D was the most affected (only 6.6% of maximal activity at 10 nM CaM). Analysis of CaM activation curves for the four Trp mutants showed that the half-maximal concentration required for activation was of 0.5 (W242V), 6.3 (W242R), 95 (W242G) and 350 nM (W242D), respectively, as compared to a value of 0.20 nM for the wild-type protein (Figure 5). As expected, the K_m for ATP of all mutants showed no significant difference when compared with that of the wild-type enzyme.

Two modified forms of adenylate cyclase (W242V and W242D) were purified to homogeneity to determine whether mutation had affected the maximum catalytic activity. The specific activity of the adenylate cyclase W242V was very close to that of the wild-type protein (2200 μ mol/min mg of protein), whereas mutant W242D had a V_m 30% lower than the wild-type enzyme. It is possible that under experimental conditions the latter mutant would not be saturated by the activator. Determination of CaM activation curves of purified enzymes gave the same K_d values as those obtained with enzyme from crude extracts.

Site-directed mutagenesis of Lys-58 and Lys-65 in recombinant adenylate cyclase

Previous experiments showed that isolated T25 fragment (residues 1–235/237) of *B. pertussis* adenylate cyclase exhibited a low but measurable enzyme activity, which indicated that it harboured the catalytic site (Ladant *et al.*, 1989). In addition, it was observed, by comparison with another CaM-activated enzyme, *Bacillus anthracis* adenylate cyclase (Escuyer *et al.*, 1988), that a 17 amino acid polypeptide was similar in both proteins and contained the sequence G---GKS (AKS in *B. pertussis*) (Figure 6) which is known to be part of many ATP-binding proteins (Walker *et al.*, 1982; Miller and Amons, 1985; Fry *et al.*,

1986). This suggested that the nucleotide-binding site might be located near Lys-65. In order to substantiate this hypothesis we substituted Gln for Lys-65. Another basic residue, Lys-58, situated in the vicinity of the former Lys residue was also submitted to the same substitution. Both mutants (K65Q and K58Q) showed drastically reduced catalytic activity in crude extracts as compared to wild-type protein at saturating or near-saturating concentrations of CaM (Table II). As a first approximation we can assume that both mutants were much less, or not at all, affected in their CaM-binding properties. Pure K65Q mutant adenylate cyclase exhibited a specific activity of 2.2 μ mol/min mg of protein when assayed with 2 mM ATP and 100 nM CaM. Determination of enzyme activity as a function of ATP and CaM concentration indicated a K_m for ATP of 1.5 mM and half-maximum concentration required for activation by CaM of 0.5 nM. The latter value is most probably an overestimate since enzyme concentration in the assay mixture was of the same order of magnitude to ensure accurate determination of its activity. As in the case of wild-type or W242 mutants, the adenylate cyclase K65Q was converted by proteolysis with trypsin into complementary T25 and T18 peptides in the presence of CaM (data not shown).

Discussion

Although adenylate cyclases from various organisms have been the subject of intense study for the past 15 years not much is known about the detailed molecular structure of their catalytic or regulatory centres. The genes coding for adenylate cyclase from two very different bacteria, *B. pertussis* (a Gram-negative organism) and *B. anthracis* (a Gram-positive bacillus), were cloned and sequenced (Escuyer *et al.*, 1988; Glaser *et al.*, 1988a; Mock *et al.*, 1988). We observed that these enzymes although having similar properties, displayed only limited sequence similarity, except in three regions. The first region (situated between residues 54–70 in *B. pertussis* enzyme and between residues 342–358 in *B. anthracis* adenylate cyclase) contained a sequence that resembled the generally accepted binding site for ATP (Walker *et al.*, 1982; Miller and Amons, 1985; Fry *et al.*, 1986) (Figure 6). We have demonstrated here in the case of *B. pertussis* adenylate cyclase that, indeed, this region must be closely involved in the enzyme active centre. From amino acid sequence comparison we can therefore confidently infer that the cognate site is also involved in the catalytic site of *B. anthracis*. The data presented are also in line with biochemical experiments using the pure enzyme fragments and suggesting that the catalytic centre was located in the first half of the protein (residues 1–235/237) secreted in *B. pertussis* culture supernatants (Ladant *et al.*, 1989).

The case of the CaM regulator site is more complex. When analysing the amino acid sequence of the protein, as predicted from the nucleotide sequence, we were prompted to suggest a tryptophan-containing peptide (Trp-242) might be involved in CaM-binding as it is in some known cases (Glaser *et al.*, 1988a). However, comparison with the *B. anthracis* enzyme did not display any significant primary structure similarity in the corresponding region. It was observed, however, that the region was bracketed by two regions of strong similarity (Escuyer *et al.*, 1988). This suggested either that we are in the presence of regions flanking a polypeptide fold where CaM is accommodated, and/or that one of these regions

contacts CaM while the other one would be involved in the information transfer between the CaM-binding site and the catalytic centre. In any case the fact that alteration of Trp-242 results in a much lower CaM-binding demonstrates that this region must be somehow involved in the association between the two proteins. The fact that there is no conservation of the corresponding primary structure indicates that it is a secondary rather than primary structure which is important for binding. It should be emphasized here that our results are the first experimental data where the putative α -helix involved in CaM-binding has been modified, and thus shown to influence binding.

In keeping with this hypothesis are the results of the specific mutation effects we observed. Aspartate, when replacing Trp diminishes drastically CaM interaction; it is known that this amino acid residue, perhaps through its specific interaction with water molecules, is often a helix breaker, whereas aromatic amino acid residues (namely Trp and Phe), are helix formers. A branched-chain amino acid such as Val should be, in this context, more like Trp and Gly more like Asp, as they indeed are. Thus we propose, in line with Cox *et al.* (1985) that an α -helix is part of the CaM-binding site of bacterial pathogen adenylate cyclase, as is the case for other CaM-binding proteins. This could correspond to a dipolar moment orienting CaM in a correct position with respect to the enzyme; the conserved sites would in such a picture be more likely to correspond to zones of contact between the enzyme and CaM.

Materials and methods

Chemicals

Adenine nucleotides, restriction enzymes and T₄ DNA ligase were from Boehringer Mannheim. L-1-tosylamido-2-phenylethyl chloromethyl ketone (TPCK) trypsin, soybean trypsin inhibitor, bovine brain CaM and CaM-agarose were from Sigma. Oligonucleotides were synthesized according to the phosphoramidate method using a commercial DNA synthesizer (Applied Biosystems). [α -³²P]ATP (3000 Ci/mmol), [³H]cAMP (40 Ci/mmol) and Na¹²⁵I (1000 Ci/mmol) were obtained from the Radiochemical Centre (Amersham, UK).

Bacterial strain and growth conditions

E. coli strain BMH7118 was used in this work (Gronenbrun, 1976). This strain expresses a *lacI*^q gene and a deleted *lacZ* gene, *lacZ*M15. The product of the first gene represses the lactose promoter in the absence of the inducer (isopropyl- β -D-thiogalactoside) (IPTG), whereas the product of the second gene, M15, is an acceptor in α -complementation (Ullmann and Perrin, 1970). Cultures were performed in LB medium (Miller, 1972) supplemented with 100 μ g/ml ampicillin, and when necessary 0.1 mM IPTG until the end of the exponential phase of growth. Bacteria were harvested by centrifugation for 30 min at 5000 g, then disrupted by sonication (2 \times 3 min at 20 KHz and 100 watts) after being resuspended in 20 mM K-phosphate buffer (pH 7.4).

Plasmids

Plasmid pDIA5202 harbours a truncated *cya* gene, fused in frame after the 459th codon with the α -*lacZ*' gene (Figure 1). The fusion gene encodes a 604 residue long protein. The fusion to the α segment of β -galactosidase increased the stability of the adenylate cyclase, allowing a simplified purification of the protein.

Site-directed mutagenesis and sequence analysis

Oligonucleotide directed mutagenesis was performed using the Amersham kit following the supplier's instructions. The Lys (AAA) codon at position 58 was modified to a Gln (CAA) codon using the oligonucleotide: GTGGCCACCC~~AA~~GGATTGG. The Lys (AAG) codon at position 65 was modified to a Gln (CAG) codon using the oligonucleotide GTGCACGCC~~AG~~TCGTCCG. Both modifications were performed on the *Bam*HI-*Eco*RV DNA fragment cloned in phage M13tg130 (Nakamaye and Eckstein, 1986). The Trp (UUG) codon at position 242

was first modified to a Asp (GAC) codon using the oligonucleotide CTTGTTGG~~AC~~AAAATCGC. This Asp codon was then modified to a Val (GTC) codon using the oligonucleotide CTTGTTGG~~TC~~AAAATCGC. Finally the Trp codon was also modified in a single step to either Gly (GGG), or Arg (CGG) codons using the ambiguous oligonucleotide GACTTGTG(C/G)GGAAAATCGC. The modifications at position 242 were performed on the *Eco*RV-*Eco*RI DNA fragment cloned in phage M13tg130. For each mutagenesis the mutated fragment was controlled for the absence of any other mutations by the dideoxynucleotide sequencing method (Sanger *et al.*, 1977).

Analytical procedures

Adenylate cyclase activity was measured as previously described (Ladant *et al.*, 1986) in 100 μ l of medium containing 50 mM Tris-HCl (pH 8), 6 mM MgCl₂, 0.1 mM CaCl₂ (or 2 mM EGTA), bovine brain CaM (between 0.03 nM and 10 μ M), 0.1 mM [³H]cAMP (10⁴ c.p.m./assay) and 2 mM [α -³²P]ATP (5 \times 10⁵ c.p.m./assay). One unit of adenylate cyclase activity corresponds to 1 μ mol of cAMP formed in 1 min at 30°C and pH 8. β -galactosidase activity was measured according to Pardee *et al.* (1959). Iodination of adenylate cyclase (0.08–0.25 mol of [¹²⁵I]mol of enzyme) and cleavage of protein by trypsin were performed exactly as previously described (Ladant *et al.*, 1988). Cleaved products were analysed by SDS-PAGE as described by Laemmli (1970) and detected by exposure of dried gels at -80°C to Kodak X-O-Mat AR films for 4–48 h with intensifying screens.

Protein concentration was measured according to Bradford (1976) or by amino acid analysis on a Biotronik amino acid analyzer LC 5001 using a single column procedure. Samples containing 10–15 μ g of protein were hydrolysed in vacuum in 0.1 ml of 6 M HCl for 20 h at 110°C.

Acknowledgements

We thank A. Ullmann for constructive comments on the manuscript, A.-M. Gilles for amino acid analysis, S. Michelson for useful suggestions, M. Mock for helpful discussion and M. Ferrand for expert secretarial help. We thank also the students participating at the 'Cours de Microbiologie', Institut Pasteur, for their confirmatory contribution to the kinetic analysis of *B. pertussis* adenylate cyclase mutants. This work was supported by grants URA 150 (UA 1129) and 955710/1.88 from the Centre National de la Recherche Scientifique.

References

- Blumenthal, D.K., Takio, K., Edelman, A.M., Charbonneau, H., Titani, K., Walsh, K.A. and Krebs, E.G. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 3187–3191.
- Bradford, M.M. (1976) *Anal. Biochem.*, **72**, 248–254.
- Buschmeier, B., Meyer, H.E. and Mayr, G.W. (1987) *J. Biol. Chem.*, **262**, 9454–9462.
- Confer, D.L. and Eaton, J.W. (1982) *Science*, **217**, 948–950.
- Cox, J.A., Comte, M., Fitton, J.E. and DeGrado, W.F. (1985) *J. Biol. Chem.*, **260**, 2527–2534.
- Escuyer, V., Duflo, E., Sezer, O., Danchin, A. and Mock, M. (1988) *Gene*, **71**, 293–298.
- Fry, D.C., Kuby, S.A. and Mildvan, A.S. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 907–911.
- Glaser, P., Ladant, D., Sezer, O., Pichot, F., Ullmann, A. and Danchin, A. (1988a) *Mol. Microbiol.*, **2**, 19–30.
- Glaser, P., Sakamoto, H., Bellalou, J., Ullmann, A. and Danchin, A. (1988b) *EMBO J.*, **7**, 3997–4004.
- Greenlee, D.U., Andreassen, T.J. and Storm, D.R. (1982) *Biochemistry*, **21**, 2759–2764.
- Gronenbrun, B. (1976) *Mol. Gen. Genet.*, **148**, 243–250.
- Hanley, R.M., Means, A.R., Kemp, B.E. and Shenolikar, S. (1988) *Biochem. Biophys. Res. Commun.*, **152**, 122–128.
- Hanski, E. and Farfel, Z. (1985) *J. Biol. Chem.*, **260**, 5526–5532.
- James, P., Maeda, M., Fischer, R., Verma, A.K., Krebs, J., Penniston, J.T. and Carafoli, E. (1988) *J. Biol. Chem.*, **263**, 2905–2910.
- Kennelly, P.J., Edelman, A.M., Blumenthal, D.K. and Krebs, E.G. (1987) *J. Biol. Chem.*, **262**, 11958–11963.
- Ladant, D. (1988) *J. Biol. Chem.*, **263**, 2612–2618.
- Ladant, D., Brézin, C., Alonso, J.-M., Crenon, I. and Guiso, N. (1986) *J. Biol. Chem.*, **261**, 16264–16269.
- Ladant, D., Michelson, S., Sarfati, R., Gilles, A.-M., Predeleanu, R. and Bărzău, O. (1989) *J. Biol. Chem.*, **264**, in press.
- Laemmli, U.K. (1970) *Nature*, **227**, 680–685.
- Miller, J.H. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor

- Laboratory Press, Cold Spring Harbor, NY.
- Miller, W. and Amons, R. (1985) *FEBS Lett.*, 186, 1-7.
- Mock, M., Labruyère, D., Glaser, P., Danchin, A. and Ullmann, A. (1988) *Gene*, 64, 277-284.
- Monneron, A., Ladant, D., d'Alayer, J., Bellalou, J., Bârzu, O. and Ullmann, A. (1988) *Biochemistry*, 27, 536-539.
- Nakamaye, K.L. and Eckstein, F. (1986) *Nucleic Acids Res.*, 14, 9676-9698.
- O'Neil, K.T., Wolfe, H.R., Jr, Erickson-Viitanen, S. and DeGrado, W.F. (1987) *Science*, 236, 1454-1456.
- Pardee, A.B., Jacob, F. and Monod, J. (1959) *J. Mol. Biol.*, 1, 165-168.
- Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA*, 74, 5463-5467.
- Shattuck, R.L., Oldenburg, D.J. and Storm, D.R. (1985) *Biochemistry*, 24, 6358-6362.
- Ullmann, A. and Perrin, D. (1970) *The Lactose Operon*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Walker, J.E., Saraste, M., Runswick, M.J. and Gay, N.J. (1982) *EMBO J.*, 1, 945-951.
- Weiss, A.A. and Hewlett, E.L. (1986) *Annu. Rev. Microbiol.*, 40, 661-686.
- Wolff, J., Cook, J.H., Goldhammer, A.R. and Berkowitz, S.A. (1980) *Proc. Natl. Acad. Sci. USA*, 77, 3841-3844.

Received on December 12, 1988; accepted on January 3, 1989

XP-002100623

P.D. 1998
p. 5752-5756 = 5

A bacterial two-hybrid system based on a reconstituted signal transduction pathway

GOUZEL KARIMOVA, JOSETTE PIDOUX, AGNES ULLMANN, AND DANIEL LADANT*

Unité de Biochimie Cellulaire (Centre National de la Recherche Scientifique, Unité de Recherche Associée 1129), Institut Pasteur, 75724 Paris Cedex 15, France

Communicated by Boris Magasanik, Massachusetts Institute of Technology, Cambridge, MA, March 2, 1998 (received for review December 2, 1997)

ABSTRACT We describe a bacterial two-hybrid system that allows an easy *in vivo* screening and selection of functional interactions between two proteins. This genetic test is based on the reconstitution, in an *Escherichia coli* *cya* strain, of a signal transduction pathway that takes advantage of the positive control exerted by cAMP. Two putative interacting proteins are genetically fused to two complementary fragments, T25 and T18, that constitute the catalytic domain of *Bordetella pertussis* adenylate cyclase. Association of the two-hybrid proteins results in functional complementation between T25 and T18 fragments and leads to cAMP synthesis. Cyclic AMP then triggers transcriptional activation of catabolic operons, such as lactose or maltose, that yield a characteristic phenotype. In this genetic test, the involvement of a signaling cascade offers the unique property that association between the hybrid proteins can be spatially separated from the transcriptional activation readout. This permits a versatile design of screening procedures either for ligands that bind to a given "bait," as in the classical yeast two-hybrid system, or for molecules or mutations that block a given interaction between two proteins of interest.

Most biological processes involve specific protein-protein interactions. General methodologies to identify interacting proteins or to study these interactions have been developed extensively. Among them, the yeast two-hybrid system currently represents the most powerful *in vivo* approach to screen for polypeptides that could bind to a given target protein. Originally developed by Fields and coworkers (1, 2), it utilizes hybrid genes to detect protein-protein interactions by means of activation of a reporter-gene expression (3, 4). In essence, the two putative protein partners are genetically fused to the DNA-binding domain of a transcription factor and to a transcriptional activation domain, respectively. A productive interaction between the two proteins of interest will bring the transcriptional activation domain in the proximity of the DNA-binding domain and will trigger the transcription of an adjacent reporter gene (usually *lacZ* or a nutritional marker), giving a screenable phenotype. Very recently, Rossi *et al.* (5) described a different approach, a mammalian "two-hybrid" system, which uses β -galactosidase complementation to monitor protein-protein interactions in intact eukaryotic cells.

A bacterial equivalent of the two-hybrid system has not yet been reported. Phage display (6, 7) and double-tagging assay (8) represent alternative approaches to screen complex libraries of proteins for direct interaction with a given ligand. However, these techniques do not allow an *in vivo* selection of the relevant clones.

We describe here a two-hybrid system in *Escherichia coli* in which the proteins of interest are genetically fused to two complementary fragments of the catalytic domain of *Borde-*

tella pertussis adenylate cyclase (9, 10). Interaction between the two proteins results in functional complementation between the two adenylate cyclase fragments leading to cAMP synthesis, which, in turn, can trigger the expression of several resident genes. Using this assay, one can select specific clones expressing a protein that interacts with a given target, by a simple genetic screening.

MATERIALS AND METHODS

Strain and Growth Media. DHP1 is an adenylate cyclase-deficient (*cya*) derivative of DH1 (F^- , *glnV44(AS)*, *recA1*, *endA1*, *gyrA96 (Nal^r)*, *thi1*, *hsdR17*, *spoT1*, *rfaD1*) (25) and was isolated by using phosphomycin as a selection antibiotic (11). Growth media used were the rich Luria-Bertani (LB) medium or the synthetic medium M63 (12) supplemented with 1% carbon source. Antibiotic concentrations were as follows: 100 μ g/ml ampicillin and 30 μ g/ml chloramphenicol. Screening for the ability to ferment sugars was performed either on MacConkey agar plates containing 1% maltose, or on LB plates containing 40 μ g/ml X-gal (5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside) and 0.5 mM isopropyl- β -D-thiogalactopyranoside.

Plasmids. The plasmid pCm-AHL1 is an expression vector for the full catalytic domain of adenylate cyclase. It was generated by subcloning a 1.5-kb *PvuII* fragment from pACTAC1322 (13) containing the first 384 codons of *cyaA* under the control of the *lac UV5* promoter [cAMP/catabolite gene activator protein (CAP)-independent], into pACYC184 linearized by *EcoRV-HincII*. Plasmid pT25 is a derivative of pACYC184 that encodes the T25 fragment of CyaA (amino acids 1-224) in frame with a multicloning site sequence, under the control of the *lac UV5* promoter. Plasmid pT18 is a derivative of pBluescript II KS (Stratagene), compatible with pT25, that encodes the T18 fragment of CyaA (amino acids 225-399) in frame with the multicloning site sequence of pBluescript II KS. The leucine zipper region of the yeast protein GCN4 was amplified by PCR from plasmid pHB16 (14) (a gift from H. Bedouelle, Institut Pasteur) using primers: lz1p1 (CTGCAGGTACCTATCCAGCGTATGAAA) and lz1p2 (TGAGGGTACCCACGTTCCACCACCAG). The amplified sequence was cleaved by *KpnI* and cloned into the *KpnI* site of pT25 and pT18 to yield plasmids pT25-*zip* and pT18-*zip*, which encode, respectively, the T25 and T18 fragments fused in frame with this 35-aa-long leucine zipper. Plasmids pT25-Tyr and pT18-Tyr express, respectively, the T25 and T18 fragments fused in frame with the first 302 aa of the dimeric tyrosyl-tRNA synthetase from *Bacillus stearothermophilus* (15). The corresponding part of the *tyrRS* gene was amplified by PCR from plasmid M13EL (15) (a gift from H. Bedouelle) by using primers *tyrS1* (AGAGGTACCGGACATGGATTGCT) and *tyrS2* (GCCGGTACCGCCGCTGTCAA-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

© 1998 by The National Academy of Sciences 0027-8424/98/955752-5\$2.00/0
PNAS is available online at <http://www.pnas.org>.

Abbreviations: LB, Luria-Bertani; X-gal, 5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside; CAP, catabolite gene activator protein.

*To whom reprint requests should be addressed at: Unité de Biochimie Cellulaire, Institut Pasteur, 28, rue du Docteur Roux, 75724 Paris Cedex 15, France. e-mail: ladant@pasteur.fr.

ATTGGC), cleaved by *KpnI* and cloned into the *KpnI* site of pT25 and pT18. The plasmid pT25-prp11 that expresses the T25 fused to the yeast-splicing factor Prp11 (16) was constructed as follows. First, we constructed pT25-2, a derivative of pT25 by inserting the sequence GCCCGGGG between the *PstI* and *BamHI* site of pT25, to change the reading frame of the *BamHI* site. Then, a 0.9-kb *BamHI* fragment of plasmid pPL253 (a gift from P. Legrain, Institut Pasteur) encompassing the *prp11* gene was subcloned into the corresponding site of pT25-2. To construct plasmid pT18-prp21, expressing T18 fused to the yeast splicing factor Prp21 (16), the gene encoding the Prp21 protein was amplified by PCR from plasmid pPL182 (a gift from P. Legrain) by using primers CCCGGTACCGA-TGGAACCAGAAGATAC and CCCGGTACCGAGTTT-TACTTTTCTTTAACC, cleaved by *KpnI* and cloned into the *KpnI* site of pT18.

Analytical Methods. β -galactosidase assays were performed on toluenized bacterial suspensions, as described in Pardee et al. (17). One unit of activity corresponds to 1 nmol of *o*-nitrophenyl β -D-galactoside hydrolyzed per min at 28°C.

Bacterial extracts were prepared from cells grown overnight at 30°C in LB medium supplemented with ampicillin and chloramphenicol. Pelleted cells were resuspended in 8 M urea/20 mM Hepes-Na, pH 7.5, and disrupted by sonication. Adenylate cyclase assays were performed on extracts as described previously (9) in the presence of 1 μ M of calmodulin. One unit of activity corresponds to 1 nmol of cAMP formed per min at 37°C. Cyclic AMP measurements were done by an ELISA assay. Briefly, a cAMP-biotinylated-BSA conjugate was coated on ELISA plates, and nonspecific protein-binding sites were blocked with BSA. Boiled bacterial cultures were

then added, followed by diluted rabbit anti-cAMP antiserum in 50 mM Hepes, pH 7.5/150 mM NaCl/0.1% Tween 20 (HBST buffer) containing 10 mg/ml BSA. After overnight incubation at 4°C, the plates were washed extensively with HBST, then goat anti-rabbit IgG coupled to alkaline phosphatase (AP) was added and incubated for 1 hr at 30°C. After washing, the AP activity was revealed by 5'-para-nitrophenyl phosphate. cAMP concentrations were calculated from a standard curve established with known concentrations of cAMP diluted in LB medium.

RESULTS

Principle. *B. pertussis* produces a calmodulin-dependent adenylate cyclase toxin encoded by the *cyaA* gene (18–20). The catalytic domain is located within the first 400 aa of this 1,706-residue-long protein (10, 19). It exhibits a high catalytic activity ($k_{cat} = 2,000 \text{ s}^{-1}$) in the presence of calmodulin (CaM), and a low but detectable activity ($k_{cat} = 2 \text{ s}^{-1}$) in the absence of this activator (9, 21). Biochemical studies revealed that the catalytic domain can be proteolytically cleaved into two complementary fragments, T25 and T18, which remain associated in the presence of CaM in a fully active ternary complex (9, 10, 22). In the absence of CaM, the mixture of the two fragments did not exhibit detectable activity, suggesting that the two fragments are not able to reassociate to yield basal CaM-independent activity. We reasoned that when expressed in an adenylate cyclase-deficient *E. coli* strain (*E. coli* lacks CaM or CaM-related proteins), the T25 and T18 fragments fused to putative interacting proteins would reassociate and lead to cAMP synthesis (Fig. 1).

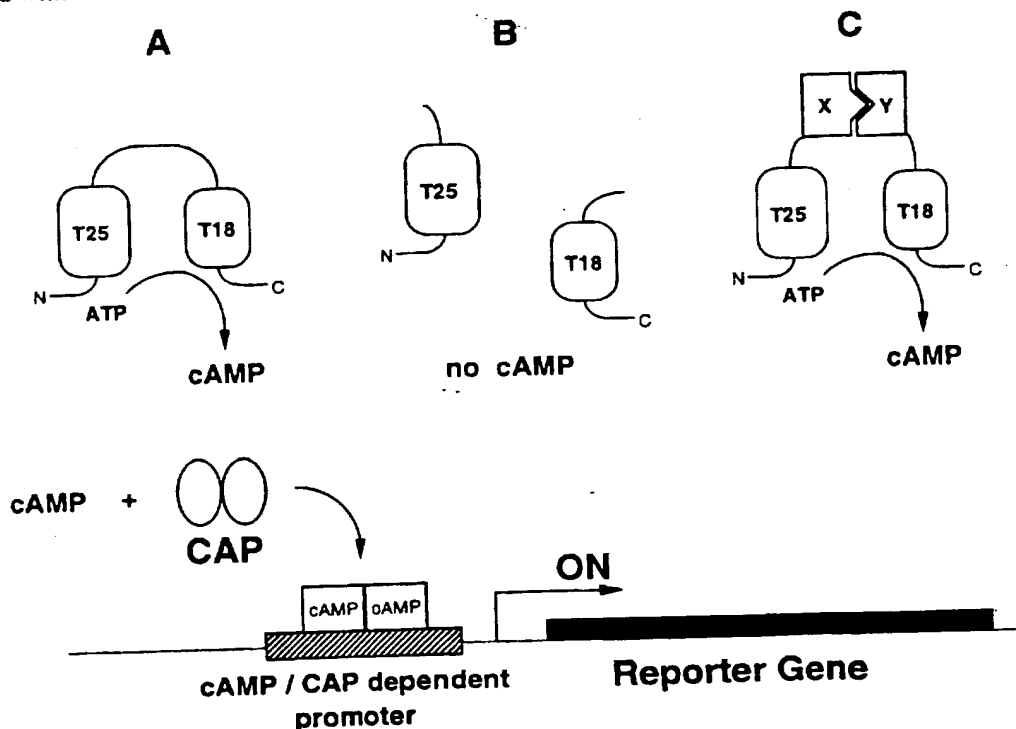


FIG. 1. Principle of an *E. coli* two-hybrid system based on functional complementation of CyaA fragments. (Upper) Schematic of the basic principle of *in vivo* complementation between the two fragments of the catalytic domain of *B. pertussis* adenylate cyclase. The two boxes represent the T25 and T18 fragments corresponding to amino acids 1–224 and 225–399 of the CyaA protein. In A, the full-length catalytic domain (residues 1–399), when expressed in *E. coli*, exhibits a basal calmodulin-independent activity that results in cAMP synthesis. In B, the two fragments, T25 and T18, when coexpressed as independent polypeptides, are unable to interact and no cAMP synthesis occurs. In C, the two fragments, fused to two interacting proteins, X and Y, are brought into close proximity, resulting in functional complementation followed by cAMP production. (Lower) Schematic of the readout of the complementation. cAMP, synthesized in an *E. coli* *cya* strain by the complementing T25 and T18 pairs, binds to the catabolite gene activator protein, CAP. The cAMP/CAP complex then can recognize specific promoters and switch on the transcription of the corresponding genes. These reporter genes can be either natural *E. coli* genes, such as *lacZ* or *mal* genes, or synthetic ones, such as antibiotic-resistance genes fused to a cAMP/CAP-dependent promoter.

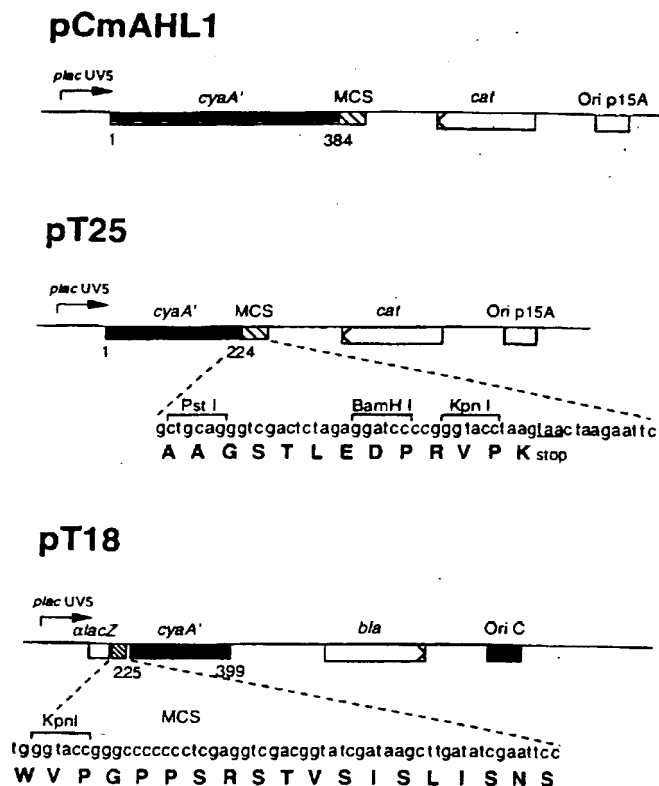


FIG. 2. Schematic representation of plasmids. The open boxes represent the ORFs of β -lactamase (*bla*) and chloramphenicol acetyl transferase (*cat*) genes. The solid boxes correspond to the ORF of *cyaA'*, with codon numbers indicated below. The hatched boxes correspond to the multicloning site sequences (MCS) that are fused at the indicated position of the *cyaA* ORF. The origin of replication of the plasmids is indicated by shaded boxes.

Functional analysis of *B. pertussis* adenylate cyclase activity can be easily monitored in an *E. coli* strain deficient in endogenous adenylate cyclase. In *E. coli*, cAMP bound to the transcriptional activator, CAP, is a pleiotropic regulator of the expression of various genes, including genes involved in the catabolism of carbohydrates such as lactose or maltose (23). Hence, *E. coli* strains lacking cAMP are unable to ferment lactose or maltose. When the entire catalytic domain of CyaA (amino acids 1–399) is expressed in *E. coli* *lacZ* (plasmid pDIA5240), its calmodulin-independent residual activity is sufficient to complement an adenylate cyclase-deficient strain and to restore its ability to ferment lactose or maltose (24). This can be scored either on indicator plates (i.e., LB-X-gal or MacConkey media supplemented with maltose) or on selective

media (minimal media supplemented with lactose or maltose as unique carbon source).

Design of a Two-Hybrid System Based on Functional Complementation Between T25 and T18 Fragments of CyaA. We first constructed two compatible plasmids (derived from pACYC184 and pBluescript II KS) that express either the T25 fragment corresponding to amino acids 1–224 of CyaA or the T18 fragment corresponding to amino acids 225–399. A multicloning site was fused to the C-terminal end of T25 to facilitate construction of fusions with foreign proteins. Similarly, the T18 fragment was fused in frame to *lacZ* of pBluescript II KS, downstream of its multicloning site (Fig. 2).

The two plasmids, pT25 and pT18, were cotransformed in DHP1, a *cya* derivative of the *E. coli* strain DH1 (25), and plated on MacConkey agar supplemented with maltose. As expected, no spontaneous complementation between the two isolated (independently expressed) fragments could be detected *in vivo*: all the transformants were white (see Table 1). When the DHP1 strain was transformed with a plasmid expressing the full catalytic domain, all colonies were red (Table 1).

To test whether functional complementation between T25 and T18 could be brought about by fusing them to interacting proteins, we inserted, within the multicloning site of both pT25 and pT18, a DNA sequence that codes for a 35-aa-long leucine zipper derived from protein GCN4, a yeast transcriptional activator (14). When the resulting plasmids, pT25-zip and pT18-zip were cotransformed in DHP1 and plated on MacConkey/maltose media, the resulting colonies became red after 24–30 hr of growth at 30°C (Table 1).

Control experiments were carried out in which pT25-zip was cotransformed with pT18 or pT18-zip was cotransformed with pT25. None of the transformants exhibited complementation, demonstrating that the functional complementation of T25-zip and T18-zip was mediated by the interaction of their leucine zipper motif. The efficiency of complementation could be further quantified by measuring in liquid cultures either cAMP levels or β -galactosidase activities (Table 1).

Adenylate cyclase activities of the different transformants were measured in cell extracts in the presence of CaM that binds tightly to T25 and T18 fragments to form the active adenylate cyclase complex. As shown in Table 1, only the extract from DHP1/pT25-zip/pT18-zip exhibited a significant enzymatic activity. In the absence of CaM, adenylate cyclase activity still could be detected in the extract of DHP1/pT25-zip/pT18-zip, indicating that the noncovalent association of T25-zip and T18-zip, mediated by their leucine zipper moiety, was able to restore the basal enzymatic activity, which was sufficient to sustain *in vivo* cAMP synthesis (Table 1). The lack of adenylate cyclase activity in the extracts of the three other types of transformants (Table 1) indicates that at least one of the two complementary fragments of adenylate cyclase was missing, most probably as a consequence of its *in vivo* proteolytic degradation. Therefore, it would appear that the associ-

Table 1. Analysis of complementation in DHP1 strain

Plasmids	Phenotype on MacConkey/maltose	β -Galactosidase, units/mg dry weight bacteria	cAMP, pmol/mg dry weight bacteria	Adenylate cyclase activity*	
				+CaM	-CaM
None	White	179	<10	<1	<0.01
pCm-AHL1	Red/24 hr	6,650	3,400	13,000	10
pT25 + pT18	White/72 hr	130	<10	<1	<0.01
pT25 + pT18-zip	White/72 hr	183	<10	<1	<0.01
pT25-zip + pT18	White/72 hr	178	<10	<1	<0.01
pT25-zip + pT18-zip	Red/30 hr	4,750	1,100	10,000	4

Bacteria were grown in LB at 30°C in the presence of 0.5 mM isopropyl β -D-thiogalactoside plus appropriate antibiotics. The results represent the average values obtained for at least five independent cultures, which differed by less than 10%.

*nmol cAMP/min per mg protein; when present in the assays, CaM was at a concentration of 1 μ M.

Table 2. Complementation between various chimeric proteins

Plasmids	Phenotype on MacConkey/maltose	β -Galactosidase, units/mg dry weight bacteria	cAMP, pmol/mg dry weight bacteria
pT25-Tyr + pT18-Tyr	Red/40 hr	2,800	580
pT25-Tyr + pT18	White/96 hr	193	<10
pT25 + pT18-Tyr	White/96 hr	183	<10
pT25-Tyr + pT18-zip	White/96 hr	134	<10
pT25-zip + pT18-Tyr	White/96 hr	126	<10
pT25-prp11 + pT18-prp21	Red/40 hr	850	65

Bacteria were grown in LB at 30°C in the presence of 0.5 mM isopropyl β -D-thiogalactoside plus appropriate antibiotics. The results represent the average values obtained for at least five independent cultures.

ation of T25-zip and T18-zip, through their leucine zipper motif, not only resulted in their functional complementation but also in their stabilization. Stabilization of protein fragments (α and ω peptides) through complementation (26) has also been observed for β -galactosidase (A.U., unpublished results).

Screening for *in Vivo* Protein-Protein Interactions by Using Functional Complementation of T25 and T18. We then examined whether the complementation between T25 and T18 could be used to analyze interactions between proteins larger than the 35-residue-long leucine zipper motif. A DNA fragment that encodes the N-terminal part (residues 1–302) of the dimeric tyrosyl tRNA synthetase from *Bacillus stearothermophilus* (15) was subcloned into the multicloning site of plasmids pT25 and pT18. The resulting plasmids, pT25-TyrRS and pT18-TyrRS, when cotransformed in DHP1, yielded red transformants on MacConkey/maltose. The transformants synthesized cAMP and expressed β -galactosidase (Table 2). Control transformations confirmed that the TyrRS moiety was responsible for the functional complementation between T25-TyrRS and T18-TyrRS (Table 2). Furthermore, as shown in Table 2, no complementation occurred when T25-TyrRS was cotransformed with pT18-zip or vice versa. This demonstrates that the complementation was dictated by the specificity of recognition of the polypeptides fused to the two fragments, T25 and T18. It also demonstrates that functional complementation and stabilization of the chimeric proteins occur only upon specific interactions between the two partners.

We further showed (Table 2) that the bacterial two-hybrid system could detect interaction between the yeast-splicing factors Prp11 and Prp21 (fused to T25 and T18, respectively) that was characterized previously in the yeast two-hybrid assay (16). This demonstrates that this bacterial complementation assay can reveal association between eukaryotic proteins.

To mimic a screening procedure we mixed plasmids pT18-zip and pT18-TyrRS with about a 5-fold excess of pT18 and cotransformed this mixture in DHP1 with either pT25 or

pT25-zip. The transformants were plated on LB-X-gal. All the colonies cotransformed with pT25 were white (Fig. 3). Around 20% of the colonies were blue when the cells were cotransformed with the mixture of pT18 derivatives and pT25-zip. The plasmid DNAs of these clones were further analyzed by restriction mapping. As expected, the blue colonies among the bacteria cotransformed with pT25-zip harbored only pT18-zip. In another series of experiments, pT18-zip was mixed with a 1,000-fold excess of pT18 and this mixture was transformed in DHP1 harboring pT25-zip and plated on MacConkey/maltose. Three red colonies were identified among about 3,000 white ones. Plasmid DNA analysis of the Mal^+ clones confirmed the presence of pT18-zip. Transformation of the same mixture of pT18-zip/pT18 into DHP1 harboring pT25 gave no Mal^+ clones of 10,000 analyzed (data not shown). These results indicate that the functional complementation between the adenylate cyclase fragments could be used to identify interacting proteins in *E. coli*.

Finally, we examined whether the complementation between T25 and T18 could be used in a selection procedure rather than using the screening described above. DHP1 bacteria cotransformed with complementing plasmids (pT25-zip/pT18-zip or pT25-TyrRS/pT18-TyrRS) were able to grow on minimal medium supplemented with lactose or maltose as unique carbon sources, whereas bacteria cotransformed with noncomplementing plasmids (pT25-zip/pT18-TyrRS or pT25-TyrRS/pT18-zip) did not grow (data not shown). To determine whether this selection could be used to identify interacting proteins among an excess of noninteracting ones, we performed the following "model screening" on selective media. DHP1 bacteria harboring pT25-zip and pT18-zip (expected phenotype: Lac^+) were mixed with a 10^5 excess of DHP1/pT25/pT18 (expected phenotype: Lac^-), and then 10^7 cells from this mixture were plated on minimal medium supplemented with lactose plus antibiotics. After 4–5 days at 30°C, 100–200 Lac^+ colonies appeared. Plasmid DNA analysis indicated that 18 of 20 of these colonies tested harbored pT25-zip and pT18-zip. When 10^7 DHP1/pT25/pT18 cells were plated on minimal medium/lactose, about 10 colonies were detected: these cells appeared to represent spontaneous revertants of DHP1 to a Lac^+ phenotype (because of either reversion of *cya* to *cya*⁺ or to cAMP/CAP-independent *lac* promoter mutations). This "model screening" demonstrates that bacteria expressing specific interacting proteins fused to the adenylate cyclase fragments could be selected among a large number (here a 10^5 -fold excess) of irrelevant clones.

DISCUSSION

We describe here an *E. coli* two-hybrid system that allows identification of interacting proteins by a simple genetic test (screening and/or selection). In our approach, the *in vivo* association of two interacting proteins is coupled to the production of a regulatory signaling molecule, cAMP, that in turn triggers the expression of specific reporter genes, giving rise to a selectable phenotype.

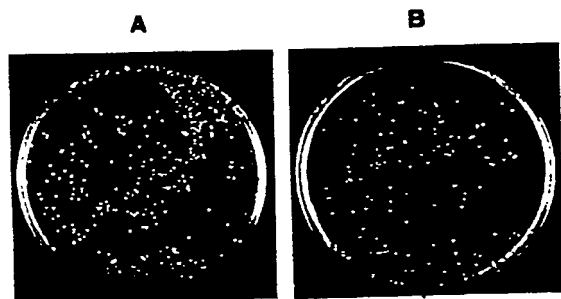


FIG. 3. Screening of interacting proteins with the bacterial two-hybrid system. DHP1 cells were cotransformed with a mixture of plasmids pT18, pT18-zip, and pT18-Tyr and either pT25 (A) or pT25-zip (B), plated on LB-X-gal agar plates containing 0.5 mM isopropyl- β -D-thiogalactopyranoside, ampicillin, and chloramphenicol, and incubated for 30 hr at 30°C. Note that the *cya*⁺ colonies are larger than the *cya*⁻ ones.

We took advantage of the modular structure of the catalytic domain of *B. pertussis* adenylate cyclase, which is composed of two complementary fragments, T25 and T18, that are both necessary to form an active enzyme, in the presence of CaM (9, 10). As shown here, the two fragments, when expressed in *E. coli* as separate entities, are unable to recognize each other and cannot reconstitute a functional enzyme. However, when T25 and T18 are fused to peptides or proteins that are able to interact, heterodimerization of these chimeric polypeptides results in a functional complementation between the adenylate cyclase fragments and, therefore, cAMP synthesis. Ultimately, cAMP, upon binding to CAP, is able to activate the transcription of catabolic operons, allowing the bacteria to ferment carbohydrates such as maltose or lactose. We have demonstrated that this bacterial two-hybrid system is able to reveal interactions between small peptides (GCN4 leucine zipper), bacterial (tyrosyl tRNA synthetase), or eukaryotic proteins (yeast Prp11/Prp21 complex).

Several characteristics of this genetic screening and/or selection provide an attractive approach to search for and analyze interacting proteins. Because our two-hybrid system involves the generation of a regulatory molecule, cAMP, the physical association of the two putative interacting proteins, can be spatially separated from the transcriptional events (activation) that are dependent on cAMP synthesis. This means that the protein-protein interaction under study does not need to take place in the vicinity of the transcription machinery as is the case for the yeast two-hybrid system. Hence, it will be possible to analyze protein interactions that occur either in the cytosol (as described here) or at the inner-membrane level. Moreover, because this genetic screen is, in essence, an assay for proximity of the fused T25 and T18 fragments, one could anticipate that this system would be particularly suitable to analyze colocalization of given proteins within multimolecular assemblies.

The readout of the complementation between T25 and T18 fragments is the transcriptional activation of cAMP/CAP-dependent genes. In this work, we took advantage of naturally occurring cAMP/CAP-dependent catabolic genes in *E. coli*. However, it is easy to design specific reporter cassettes in which any gene of interest is fused to a cAMP/CAP-dependent promoter. We currently are constructing such a system by using an antibiotic-resistance gene. This will facilitate the screening of complex libraries by a simple selection for antibiotic resistance. Alternatively, the reporter gene driven by cAMP/CAP could encode a toxic product. This could be particularly useful to search for chemical compounds or mutations that abolish a given interaction between the two studied proteins. This bacterial system is, therefore, particularly versatile as it offers the possibility of both positive and negative selections.

That this genetic test is carried out in *E. coli* greatly facilitates the screening as well as the characterization of the interacting proteins. First, it is possible to use the same plasmid constructs to screen a library to identify a putative binding partner to a given "bait," and then to express the hybrid proteins to characterize their interaction by *in vitro* binding assays. Second, the high efficiency of transformation that can be achieved in *E. coli* allows the analysis of libraries of high complexity. This is particularly useful for (i) the screening of peptides from a library made from random DNA sequences that present an affinity for a given bait protein and (ii) the exhaustive analysis of the network of interactions between the proteins of a given organism (27, 28).

In essence, our system exploits one of the fundamental principles of signal transduction, that is, signal amplification. In the system described here we relied only on the basal enzymatic activity of *B. pertussis* adenylate cyclase, i.e., in the absence of CaM, which nevertheless is sufficient to synthesize enough cAMP to activate the transcription of the *lac* and *mal* operons (we estimate, from adenylate cyclase activity measured in bacterial extracts, that about a thousand of T25 and T18 molecules per cell are expressed in the present design). This system could be rendered exquisitely sensitive by using the full catalytic potency of *B. pertussis* adenylate cyclase, i.e., in the presence of its natural activator. In this case, it is anticipated that reconstitution of only a few hybrid molecules per cell will be sufficient to elicit a detectable signal.

We thank H. Bedouelle and P. Legrain for the gift of plasmids, fruitful discussion, and comments on the manuscript. We are grateful to B. Lefèvre-Laoide for critical reading of the manuscript and to M. Berlyn for providing strain DH1. Financial support came from the Institut Pasteur and the Centre National de la Recherche Scientifique (URA 1129).

- Fields, S. & Song, O. (1989) *Nature (London)* **340**, 245-246.
- Chien, C. T., Bartel, P. L., Sternglanz, R. & Fields, S. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 9578-9582.
- Allen, J. B., Walberg, M. W., Edwards, M. C. & Elledge, S. J. (1995) *Trends Biochem. Sci.* **20**, 511-516.
- Transy, C. & Legrain, P. (1995) *Mol. Biol. Rep.* **21**, 119-127.
- Rossi, F., Charlton, C. A. & Blau, H. M. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 8405-8410.
- Smith, G. P. (1985) *Science* **228**, 1315-1317.
- Scott, J. K. & Smith, G. P. (1990) *Science* **249**, 386-390.
- Germino, F. J., Wang, Z. X. & Weissman, S. M. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 933-937.
- Ladant, D. (1988) *J. Biol. Chem.* **263**, 2612-2618.
- Ladant, D., Michelson, S., Sarfati, R. S., Gilles, A.-M., Predeleanu, R. & Bärzu, O. (1989) *J. Biol. Chem.* **264**, 4015-4020.
- Alper, M. D. & Ames, B. N. (1978) *J. Bacteriol.* **133**, 149-157.
- Miller, J. H. (1972) *Experiments in Molecular Genetics* (Cold Spring Harbor Lab. Press, Cold Spring Harbor, NY).
- Sebo, P. & Ladant, D. (1993) *Mol. Microbiol.* **9**, 999-1009.
- Blondel, A. & Bedouelle, H. (1991) *Protein Eng.* **4**, 457-461.
- Guez-Ivanier, V. & Bedouelle, H. (1996) *J. Mol. Biol.* **255**, 110-120.
- Legrain, P. & Chapon, C. (1993) *Science* **262**, 108-110.
- Pardee, A. B., Jacob, F. & Monod, J. (1959) *J. Mol. Biol.* **1**, 165-168.
- Hewlett, E. L., Urban, M. A., Manclark, C. R. & Wolff, J. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 1926-1930.
- Glaser, P., Ladant, D., Sezer, O., Pichot, F., Ullmann, A. & Danchin, A. (1988) *Mol. Microbiol.* **2**, 19-30.
- Mock, M. & Ullmann, A. (1993) *Trends Microbiol.* **1**, 187-192.
- Wolff, J., Cook, G. H., Goldhammer, A. R. & Berkowitz, S. A. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 3841-3844.
- Munier, H., Gilles, A. M., Glaser, P., Krin, E., Danchin, A., Sarfati, R. & Barzu, O. (1991) *Eur. J. Biochem.* **196**, 469-474.
- Ullmann, A. & Danchin, A. (1983) in *Advances in Cyclic Nucleotide Research* (Raven, New York), Vol. 15, pp. 1-53.
- Ladant, D., Glaser, P. & Ullmann, A. (1992) *J. Biol. Chem.* **267**, 2244-2250.
- Hanahan, D. (1983) *J. Mol. Biol.* **166**, 557-580.
- Ullmann, A., Jacob, F. & Monod, J. (1968) *J. Mol. Biol.* **32**, 1-13.
- Bartel, P. L., Roecklein, J. A., SenGupta, D. & Fields, S. (1996) *Nature Gen.* **12**, 72-77.
- Fromont, R. M., Rain, J. C. & Legrain, P. (1997) *Nature Gen.* **16**, 277-282.





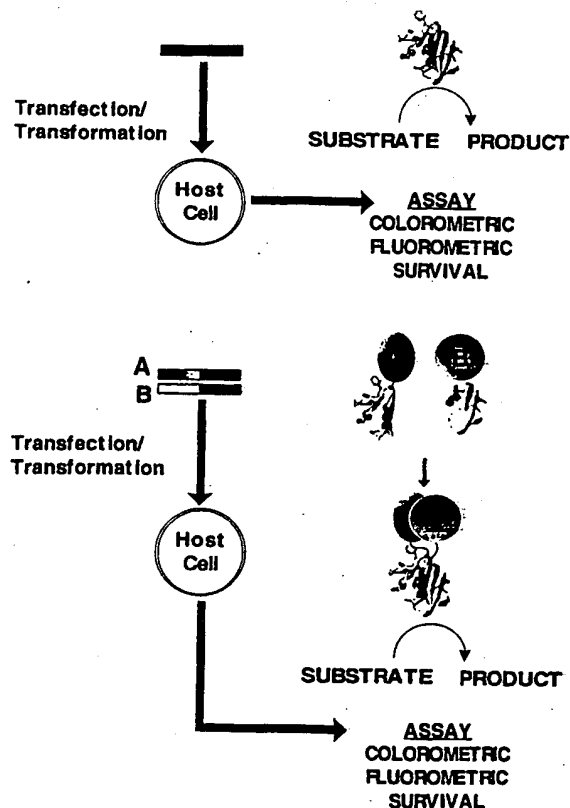
INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : G01N 33/68, 33/58, 33/573		A1	(11) International Publication Number: WO 98/34120
			(43) International Publication Date: 6 August 1998 (06.08.98)
(21) International Application Number: PCT/CA98/00068		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).	
(22) International Filing Date: 2 February 1998 (02.02.98)			
(30) Priority Data: 2,196,496 31 January 1997 (31.01.97) CA			
(71) Applicant (for all designated States except US): UNIVERSITE DE MONTREAL [CA/CA]; Postal Code 6128, Station A, Montreal, Quebec H3C 3J7 (CA).			
(72) Inventors; and (75) Inventors/Applicants (for US only): MICHNICK, Stephen, William, Watson [CA/CA]; 330 Grosvenor, Westmount, Quebec H3C 2M2 (CA). PELLETIER, Joelle, Nina [CA/CA]; Apartment #611, 10 Rosemount Avenue, Westmount, Quebec H3Y 3K4 (CA). REMY, Ingrid [CA/CA]; 8526 Berri, Montreal, Quebec H2P 2G4 (CA).			
(74) Agents: DUBUC, Jean, H. et al.; Goudreau Gage Dubuc & Martineau Walker, Stock Exchange Tower, Suite 3400, 800 Place Victoria, P.O. Box 242, Montreal, Quebec H4Z 1E9 (CA).		Published With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.	

(54) Title: PROTEIN FRAGMENT COMPLEMENTATION ASSAYS TO DETECT BIOMOLECULAR INTERACTIONS

(57) Abstract

We describe a strategy for designing and implementing protein-fragment complementation assays (PCAs) to detect biomolecular interactions *in vivo* and *in vitro*. The design, implementation and broad applications of this strategy are illustrated with a large number of enzymes with particular detail provided for the example of murine dihydrofolate reductase (DHFR). Fusion peptides consisting of N and C-terminal fragments of murine DHFR fused to GCN4 leucine zipper sequences were coexpressed in *Escherichia coli* grown in minimal medium, where the endogenous DHFR activity was inhibited with trimethoprim. Coexpression of the complementary fusion products restored colony formation. Survival only occurred when both DHFR fragments were present and contained leucine-zipper forming sequences, demonstrating that reconstitution of enzyme activity requires assistance of leucine zipper formation. DHFR fragment-interface point mutants of increasing severity (Ile to Val, Ala and Gly) resulted in a sequential increase in *E. coli* doubling times illustrating the successful DHFR fragment reassembly rather than non-specific interactions between fragments. This assay could be used to study equilibrium and kinetic aspects of molecular interactions including protein-protein, protein-DNA, protein-RNA, protein-carbohydrate and protein-small molecule interactions, for screening cDNA libraries for binding of a target protein with unknown proteins or libraries of small organic molecules for biological activity. The selection and design criteria applied here is developed for numerous examples of clonal selection, colorimetric, fluorometric and other assays based on enzymes whose products can be measured. The development of such assay systems is shown to be simple, and provides for a diverse set of protein fragment complementation applications.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

TITLE OF THE INVENTION

PROTEIN FRAGMENT COMPLEMENTATION ASSAYS TO DETECT BIOMOLECULAR INTERACTIONS

5 **FIELD OF THE INVENTION**

The present invention relates to the determination of the function of novel gene products. The invention further relates to Protein fragment Complementation Assays (PCA). PCAs allow for the detection of a wide variety of types of protein-protein, protein-RNA, protein-DNA,
10 Protein-carbohydrate or protein-small organic molecule interactions in different cellular contexts appropriate to the study of such interactions.

BACKGROUND OF THE INVENTION

Many processes in biology, including transcription,
15 translation, and metabolic or signal transduction pathways, are mediated by non-covalently-associated multienzyme complexes^{1,101}. The formation of multiprotein or protein-nucleic acid complexes produce the most efficient chemical machinery. Much of modern biological research is concerned with identifying proteins involved in cellular processes,
20 determining their functions and how, when, and where they interact with other proteins involved in specific pathways. Further, with rapid advances in genome sequencing projects there is a need to develop strategies to define "protein linkage maps", detailed inventories of protein interactions that make up functional assemblies of proteins^{2,3}. Despite the
25 importance of understanding protein assembly in biological processes, there are few convenient methods for studying protein-protein interactions *in vivo*^{4,5}. Approaches include the use of chemical crosslinking reagents

and resonance energy transfer between dye-coupled proteins^{102, 103}. A powerful and commonly used strategy, the yeast two-hybrid system, is used to identify novel protein-protein interactions and to examine the amino acid determinants of specific protein interactions^{4,6-8}. The approach
5 allows for rapid screening of a large number of clones, including cDNA libraries. Limitations of this technique include the fact that the interaction must occur in a specific context (the nucleus of *S. cerevisiae*), and generally cannot be used to distinguish induced versus constitutive interactions.

10 Recently, a novel strategy for detecting protein-protein interactions has been demonstrated by Johnsson and Varshavsky¹⁰⁸ called the ubiquitin-based split protein sensor (USPS)⁹. The strategy is based on cleavage of proteins with *N*-terminal fusions to ubiquitin by cytosolic proteases (ubiquitinases) that recognize its tertiary structure.
15 The strategy depends on the reassembly of the tertiary structure of the protein ubiquitin from complementary *N*- and *C*-terminal fragments and crucially, on the augmentation of this reassembly by oligomerization domains fused to these fragments. Reassembly is detected as specific proteolysis of the assembled product by cytosolic proteases
20 (ubiquitinases). The authors demonstrated that a fusion of a reporter protein-ubiquitin *C*-terminal fragment could also be cleaved by ubiquitinases, but only if co-expressed with an *N*-terminal fragment of ubiquitin that was complementary to the *C*-terminal fragment. The reconstitution of observable ubiquitinase activity only occurred if the *N*-
25 and *C*-terminal fragments were bound through GCN4 leucine zippers^{109,110}. The authors suggested that this "split-gene" strategy could be used as an *in vivo* assay of protein-protein interactions and analysis

of protein assembly kinetics in cells. Unfortunately, this strategy requires additional cellular factors (in this case ubiquitinases) and the detection method does not lend itself to high-throughput screening of cDNA libraries.

5 Rossi, F., C. A. Charlton, and H. M. Blau (1997) Proc. Nat. Acad. Sci. (USA) 94, 8405-8410) have reported an assay based on the classical complementation of a and w fragments of b-galactosidase (b-gal) and induction of complementation by induced oligomerization of the proteins FKBP12 and the mamalian target of rapamycin by rapamycin
10 in transfected C2C12 myoblast cell lines. Reconstitution of b-gal activity is detected using substrate fluorescein di-b-D-galactopyranoside using several fluorecence detection assays. While this assay bears some resemblance to the present invention, there are several significant distinguishing differences. First, this particular complementation
15 approach has been used for over thirty years in a vast number of applications including the detection of protein-protein interactions. Krevolin, M. and D. Kates (1993) U.S. Patent No. 5,362,625) teaches the use of this complementation to detect protein-protein interactions. Also achievement of b-gal complementation in mamalian cells has previously
20 been reported (Moosmann, P. and S. Rusconi (1996) Nucl. Acids Res. 24, 1171-1172). The individual PCAs presented here are completely *de novo* designed interaction detection assays, not described in any way previously except for publications arising from applicants laboratory. Secondly, this application describes a general strategy to develop
25 molecular interaction assays from a large number of enzyme or protein detectors, all *de novo* designed assays, whereas the b-gal assay is not

novel, nor are any general strategies or advancements over previously well documented applications given.

As in the USPS, the yeast-two hybrid strategy requires additional cellular machinery for detection that exist only in specific cellular compartments. There is therefore a need for a detection system which uses the reconstitution of a specific enzyme activity from fragments as the assay itself, without the requirement for other proteins for the detection of the activity. Preferably, the assay would involve an oligomerization-assisted complementation of fragments of monomeric or multimeric enzymes that require no other proteins for the detection of their activity. Furthermore, if the structure of an enzyme were known it would be possible to design fragments of the enzyme to ensure that the reassembled fragments would be active and to introduce mutations to alter the stringency of detection of reassembly. However, knowledge of structure is not a prerequisite to the design of complementing fragments, as will be explained below. The flexibility allowed in the design of such an approach would make it applicable to situations where other detection systems may not be suitable.

Recent advances in human genomics research has led to rapid progress in the identification of novel genes. In applications to biological and pharmaceutical research, there is now the pressing need to determine the functions of novel gene products; for example, for genes shown to be involved in disease phenotypes. It is in addressing questions of function where genomics-based pharmaceutical research becomes bogged down and there is now the need for advances in the development of simple and automatable functional assays. A first step in defining the function of a novel gene is to determine its interactions with

other gene products in an appropriate context; that is, since proteins make specific interactions with other proteins or other biopolymers as part of functional assemblies, an appropriate way to examine the function of a novel gene is to determine its physical relationships with the products of other genes.

Screening techniques for protein interactions, such as the yeast "two-hybrid" system, have transformed molecular biology, but can only be used to study specific types of constitutively interacting proteins or interactions of proteins with other molecules, in narrowly defined cellular and compartmental contexts and require a complex cellular machinery (transcription) to work. To rationally screen for protein interactions within the context of a specific problem requires more flexible approaches. Specifically, assays that meet criteria necessary not only to detecting molecular interactions, but also to validating these interactions as specific and biologically relevant.

A list of assay characteristics that meet such criteria are as follows:

- 1) Allow for the detection of protein-protein, protein-DNA/RNA or protein-drug interactions *in vivo* or *in vitro*.
- 2) Allow for the detection of these interactions in appropriate contexts, such as within a specific organism, cell type, cellular compartment, or organelle.
- 3) Allow for the detection of induced *versus* constitutive protein-protein interactions (such as by a cell growth or inhibitory factor).
- 4) To be able to distinguish specific *versus* non-specific protein-protein interactions by controlling the sensitivity of the assay.
- 5) Allow for the detection of the kinetics of protein assembly in cells.

6) Allow for screening of cDNA, small organic molecule, or DNA or RNA libraries for molecular interactions.

SUMMARY OF THE INVENTION

5 The present invention seeks to provide the above-mentioned needs for which the prior art is silent. The present invention provides a general strategy for detecting protein interactions with other biopolymers including other proteins, nucleic acids, carbohydrates or for
10 screening small molecule libraries for compounds of potential therapeutic value. In a preferred embodiment, the instant invention seeks to provide an oligomerization-assisted complementation of fragments of monomeric enzymes that require no other proteins for the detection of their activity. In one such embodiment, a protein-fragment complementation assay (PCA) based on reconstitution of dihydrofolate reductase activity by
15 complementation of defined fragments of the enzyme in *E. coli* is hereby provided. This assay requires no additional endogenous factors for detecting specific protein-protein interactions (i.e. leucine zipper interactions) and can be conveniently extended to screening cDNA, nucleic acid, small molecule or protein design libraries for molecular
20 interactions. In addition, the assay can also be adapted for detection of protein interactions in any cellular context or compartment and be used to distinguish between induced *versus* constitutive protein interactions in both prokaryotic and eukaryotic systems.

25 One particular strategy for designing a protein complementation assay (PCA) is based on using the following characteristics: 1) A protein or enzyme that is relatively small and monomeric, 2) for which there is a large literature of structural and

functional information, 3) for which simple assays exist for the reconstitution of the protein or activity of the enzyme, both *in vivo* and *in vitro*, and 4) for which overexpression in eukaryotic and prokaryotic cells has been demonstrated. If these criteria are met, the structure of the enzyme is used to decide the best position in the polypeptide chain to split the gene in two, based on the following criteria: 1) The fragments should result in subdomains of continuous polypeptide; that is, the resulting fragments will not disrupt the subdomain structure of the protein, 2) the catalytic and cofactor binding sites should all be contained in one fragment, and 3) resulting new *N*- and *C*-termini should be on the same face of the protein to avoid the need for long peptide linkers and allow for studies of orientation-dependence of protein binding.

It should be understood that the above mentioned criteria do not all need to be satisfied for a proper working of the present invention. It is an advantage that the enzyme be small, preferably between 10-40 kDa. Although monomeric enzymes are preferred, multimeric enzymes can also be envisaged as within the scope of the present invention. The dimeric protein tyrosinase can be used in the instant assay. The information on the structure of the enzyme provides an additional advantage in designing the PCA, but is not necessary. Indeed, an additional strategy, to develop PCAs is presented, based on a combination of exonuclease digestion-generated protein fragments followed by directed protein evolution in application to the enzyme aminoglycoside kinase. Although the overexpression in prokaryotic cells is preferred it is not a necessity. It will be understood to the skilled artisan that the enzyme catalytic site (of the chosen enzyme) does not absolutely need to be on same molecule.

The present application explains the rationale and criteria for using a particular enzyme in a PCA. Figure 1 shows a general description of a PCA. The gene for a protein or enzyme is rationally dissected into two or more fragments. Using molecular biology techniques, the chosen fragments are subcloned, and to the 5' ends of each, proteins that either are known or thought to interact are fused. Co-transfection or transformation these DNA constructs into cells is then carried out. Reassembly of the probe protein or enzyme from its fragments is catalyzed by the binding of the test proteins to each other, and reconstitution is observed with some assay. It is crucial to understand that these assays will only work if the fused, interacting proteins catalyze the reassembly of the enzyme. That is, observation of reconstituted enzyme activity must be a measure of the interaction of the fused proteins.

A preferred embodiment of the present invention focuses on a PCA based on the enzyme dihydrofolate reductase. Expansion of the strategy to include assays in eukaryotic cells, library screening, and a specific application to problems concerning the study of integrated biochemical pathways such as signal transduction pathways, is presented. Additional assays, including those based on enzymes that can act as dominant or recessive drug selection or metabolic salvage pathways are disclosed. In addition, PCAs based on enzymes that will produce a colored or fluorescent product are also disclosed. The present invention teaches how the PCA strategy can be both generalized and automated for functional testing of novel genes, screening of natural products or compound libraries for pharmacological activity and identification of novel gene products that interact with DNA, RNA or

carbohydrates are disclosed. It also teaches how the PCA strategy can be applied to identifying natural products or small molecules from compound libraries of potential therapeutic value that can inhibit or activate such molecular interactions and how enzyme substrates and small molecule inhibitors of enzymes can be identified. Finally, it teaches how the PCA strategy can be used to perform protein engineering experiments that could lead to designed enzymes with industrial applications or peptides with biological activity.

Simple strategies to design and implement assays for detecting protein interactions *in vivo* are disclosed herein. We have designed complementary fragments of the native mDHFR that, when coexpressed in *E. coli* grown in minimal medium, allow for survival of clones expressing the two fragments, where the basal activity of the endogenous bacterial DHFR is inhibited by the competitive inhibitor trimethoprim (Fig. 3). Reconstitution of activity only occurred when both N- and C-terminal fragments of DHFR were coexpressed as C-terminal fusions to GCN4 leucine zipper sequences, indicating that reassembly of the fragments requires formation of a leucine zipper between the N- and C-terminal fusion peptides. The sequential increase in cell doubling times resulting from the destabilizing mutations directed at the assembly interface (Ile114 to Val, Ala or Gly) demonstrates that the observed cell survival under selective conditions is a result of the specific, leucine-zipper-assisted association of mDHFR fragment[1,2] with fragment[3], as opposed to nonspecific interactions of Z-F[3] with Z-F[1,2]. several detailed and many additional examples are given.

As demonstrated previously with the ubiquitin-based split protein sensor (USPS)⁹, a protein-fragment complementation

strategy can be used to study equilibrium and kinetic aspects of protein-protein interactions *in vivo*. The DHFR and other PCAs however, are simpler assays. They are complete systems; no additional endogenous factors are necessary and the results of complementation are observed directly, with no further manipulation. The *E. coli* cell survival assay described herein should therefore be particularly useful for screening cDNA libraries for protein-protein interactions. mDHFR expression in cells can be monitored by binding of fluorescent high-affinity substrate analogues for DHFR²⁶.

There are several further aspects of the PCAs that distinguish them from all other strategies for studying protein-protein interactions *in vivo* (except USPS). We have designed complementary fragments of enzymes that allow for controlling the stringency of the assay, and could be used to obtain estimates of the kinetics and equilibrium constants for association of two proteins. For example, with DHFR the point mutations of the wild-type enzyme Ile 114 to Val, Ala, or Gly alter the stringency of reconstitution of DHFR activity. For determining estimates of equilibrium and kinetic parameters for a specific protein-protein interaction, one could perform a series of DHFR PCA experiments with two proteins that interact with a known affinity, using the wild type or destabilizing mutant DHFR fragments. Comparison of cell growth rates in this model system with rates for a DHFR PCA using unknowns would give an estimate of the strength of the unknown interaction.

It should be understood that the present invention should not be limited to the DHFR or other PCAs presented, as it is only non-limiting embodiments of the protein complementation assay of the

present invention. Moreover, the PCAs should not be limited in the context in which they could be used. Constructs could be designed for targeting the PCA fusions to specific compartments in the cell by addition of signaling peptide sequences^{27,28}. Induced versus constitutive protein-protein interactions could be distinguished by a eukaryotic version of the PCA, in the case of an interaction that is triggered by a biochemical event. Also, the system could be adapted for use in screening for novel, induced protein-molecular associations between a target protein and an expression library.

The instant invention is also directed to a method for detecting biomolecular interactions said method comprising:

- (a) selecting an appropriate reporter molecule;
- (b) effecting fragmentation of said reporter molecule such that said fragmentation results in reversible loss of reporter function;
- (c) fusing or attaching fragments of said reporter molecule separately to other molecules; followed by
- (d) reassociation of said reporter fragments through interactions of the molecules that are fused to said fragments.

The invention also provides molecular fragment complementation assays for the detection of molecular interactions comprising a reassembly of separate fragments of a molecule, wherein reassembly of said fragments is operated by the interaction of molecular domains fused to each fragment of said molecules, and wherein reassembly of the fragments is independent of other molecular processes.

In another aspect, the present invention is directed to a method of testing biomolecular interactions comprising:

- a) generating a first fusion product comprising
- i) a first fragment of a first molecule and
 - ii) a second molecule which is different or the same as said first molecule;
- 5 b) generating a second fusion product comprising
- i) a second fragment of said first molecule; and
 - ii) a third molecule which is different from or the same as said first molecule or second molecule;
- c) allowing the first and second fusion products to
- 10 contact each other; and
- d) testing for activity regained by association of the recombined fragments of the first molecule, wherein said reassociation is mediated by interaction of the second and third molecules.
- In another novel feature, the invention is directed to a
- 15 method comprising an assay where fragments of a first molecule are fused to a second molecule and fragment association is detected by reconstitution of the first molecule's activity.
- The present invention also provides a composition comprising a product selected from the group consisting of:
- 20 (a) a first fusion product comprising:
- 1) a first fragment of a first molecule whose fragments can exhibit a detectable activity when associated and
 - 2) a second molecule that can bind (a)(1);
- (b) a second fusion product comprising
- 25 1) a second fragment of said first molecule and
- 2) a third molecule that can bind (b)(1); and
- c) both (a) and (b).

The invention further provides a composition comprising complementary fragments of a first molecule, each fused to a separate fragment of a second molecule.

5 The inventors of the present subject matter further provide a composition comprising a nucleic acid molecule coding for a fusion product, which molecule comprises sequences coding for a product selected from the group consisting of:

- (a) a first fusion product comprising:
 - 1) fragments of a first molecule whose fragments
 - 10 can exhibit a detectable activity when associated and
 - 2) a second molecule fused to the fragment of the first molecule;
- (b) a second fusion product comprising
 - 1) a second fragment of said first molecule and
 - 15 2) a second or third molecule; and
- (c) both (a) and (b).

The present invention is also directed to a method of testing for biomolecular interactions associated with: (a) complementary fragments of a first molecule whose fragments can exhibit a detectable

20 activity when associated or (b) binding of two protein-protein interacting domains from a second or third molecule, said method comprising:

- 1) creating a fusion of
 - (a) a first fragment of a first molecule whose
 - 25 fragments can exhibit a detectable activity when associated and
 - (b) a first protein-protein interacting domain;
- 2) creating a fusion of.
 - (a) a second fragment of said first molecule and

(b) a second protein-protein interacting domain that can bind said first protein-protein interacting domain;

3) allowing the fusions of (1) and (2) to contact each other; and

5 4) testing for said activity.

The instant invention further provides a composition comprising a product selected from the group consisting of:

(a) a first fusion product comprising:

10 1) a first fragment of a molecule whose fragments can exhibit a detectable activity when associated and

2) a first protein-protein interacting domain;

(b) a second fusion product comprising

1) a second fragment of said first molecule and

2) a second protein-protein interacting domain

15 that can bind said first protein-protein interacting domain; and

(c) both (a) and (b).

The invention is also directed to a composition comprising a nucleic acid molecule coding for a fusion product, which molecule comprises sequences coding for either:

20 (a) a first fusion product comprising:

1) a first fragment of a molecule whose fragments can exhibit a detectable activity when associated and

2) a first protein-protein interacting domain; or

(b) a second fusion product comprising

25 1) a second fragment of said molecule and

2) a second protein-protein interacting domain

that can bind said first protein-protein interacting domain; or

(c) both (a) and (b).

The invention also provides a method of detecting kinetics of protein assembly and screening cDNA libraries comprising performing PCA.

5 In another embodiment, the invention further provides a method of testing the ability of a compound to inhibit molecular interactions in a PCA comprising performing a PCA in the presence of said compound and correlating any inhibition with said presence.

10 In a further embodiment, the invention provides a method for detecting protein-protein interactions in living organisms and or cells, which method comprises:

(a) synthesizing probe protein fragments from an enzyme which enables dominant selection by dissecting the gene coding for the enzyme into at least two fragments;

15 (b) constructing fusion proteins with one or more molecules that are to be tested for interactions;

(c) fusing the proteins obtained in (b) with one or more of the probe fragments;

(d) coexpressing the fusion proteins; and

20 (e) detecting the reconstitution of enzyme activity.

The invention still provides a method for detecting biomolecular interactions said method comprising:

(a) selecting an appropriate reporter molecule;

(b) effecting fragmentation of said reporter molecule;

25 (c) fusing or attaching fragments of said reporter molecule separately to other molecules; followed by

(d) reassociation of said reporter fragments through interactions of the molecules that are fused to said fragments.

5 Lastly, the invention also provides a novel method of affecting gene therapy, which includes the step of providing the assays and compositions described above.

The present invention is pioneering as it is the first protein complementation assay displaying such a level of simplicity and versatility. The exemplified embodiments are protein-fragment complementation assays (PCA) based on mDHFR, where a leucine zipper directs the reconstitution of DHFR activity. Activity was detected
10 by an *E. coli* survival assay which is both practical and inexpensive. This system illustrates the use of mDHFR fragment complementation in the detection of leucine zipper dimerization and could be applied to the detection of unknown, specific protein-molecular interactions *in vivo*.

15 It should be understood that the instant invention is not limited to the PCAs presented here, as numerous other enzymes can be selected and used in accordance with the teachings of the present invention. Examples of such markers can be found in Kaufman, (1987 Genetic Eng. 9:155-198) and references found therein as well as table 1
20 of this application.

It should also be clear to the skilled artisan to which the present invention pertains that the invention is not limited to the use of leucine zippers as the two interacting molecules. Indeed, numerous other types of protein-molecule interactions can be used and identified in
25 accordance with the teaching of the present invention. The known types of motifs involved in protein-molecular interactions are well known in the art.

The present application refers to numerous prior art documents and the entire contents of all those prior art documents are herein incorporated by reference.

Other features and advantages of the present invention
5 will be apparent from the following description of the preferred embodiments thereof, the appended Examples and from the enjoined claims.

BRIEF DESCRIPTION OF THE DRAWINGS

10 FIG. 1 provides a general description of a PCA. Using molecular biology techniques, the chosen fragments of the enzyme are subcloned, and to the 5' ends of each, proteins that either are known or thought to interact are fused. Co-transfection or transformation these
15 DNA constructs into cells is then carried out and reconstitution with some assay is observed.

FIG. 2 is a scheme of the fusion constructs used in one of the embodiments of the invention. The hexahistidine peptide (6His), the homodimerizing GCN4 leucine zipper (Zipper) and mDHFR fragments (1, 2 and 3) are illustrated. The labels for the constructs are used to
20 identify both the DNA constructs and the proteins expressed from these constructs.

FIG. 3: (A) shows *E. coli* survival assay on minimal medium plates. Control: Left side of the plate: *E. coli* harboring pQE-30 (no insert); right side: *E. coli* harboring pQE-16, coding for native mDHFR.
25 Panel I: Left side of each plate: transformation with construct Z-F[1,2]; right side of each plate: transformation with construct Z-F[3]. Panel II: Cotransformation with constructs Z-F[1,2] and Z-F[3]. Panel III:

Cotransformation with constructs Control-F[1,2] and Z-F[3]. All plates contain 0.5 mg/ml trimethoprim. In panels I to III, plates on the right side contain 1mM IPTG.

(B) *E. coli* survival assay using destabilizing DHFR mutants. Panel I: Cotransformation of *E. coli* with constructs Z-F[1,2] and Z-F[3:Ile114Val]. Panel II: Cotransformation with Z-F[1,2] and Z-F[3:Ile114Ala]. Inset is a 5-fold enlargement of the right-side plate. Panel III: Cotransformation with Z-F[1,2] and Z-F[3:Ile114Gly]. All plates contain 0.5 mg/ml trimethoprim. Plates on the right side contain 1mM IPTG.

FIG. 4 features the coexpression of mDHFR fragments. (A) Agarose gel analysis of restriction pattern resulting from HincII digestion of plasmid DNA. Lane 1 contains DNA isolated from *E. coli* cotransformed with constructs Z-F[1,2] and Z-F[3]. Lanes 2 and 3 contain DNA isolated from *E. coli* transformed with, respectively, construct Z-F[3] and construct Z-F[1,2]. Fragment migration (in bp) is indicated to the right.

(B) SDS-PAGE analysis of mDHFR fragment expression. Lanes 1 to 5 show crude lysate of untransformed *E. coli* (lane 1), or *E. coli* expressing Z-F[1,2] (20.8 kDa; lane 2), Z-F[3] (18.4 kDa; lane 3), Control-F[1,2] (14.2 kDa; lane 4), and Z-F[1,2] + Z-F[3] (lane 5). Lane 6 shows 40 ml out of 2ml copurified Z-F[1,2] and Z-F[3]. Arrowheads point to the proteins of interest. Migration of molecular weight markers (in kDa) is indicated to the right.

FIG. 5 illustrates the general features of a PCA based on a survival assay such as the DHFR PCA. The assay can be used in a bacterial or a mammalian context. The inserted target DNA can be a

known sequence coding for a protein (or protein domain) of interest, or can be a cDNA library.

FIG. 6 represents an autoradiograph of a COS cell lysate after a 30 min. ^{35}S -Met-Cys pulse-labelling. The expression pattern is essentially identical to that observed in *E. coli* (see Fig. 4). The DNA transfected into the cells (or cotransfected) is indicated above the respective lanes.

FIG. 7 illustrates the results of a protein engineering application of the mDHFR bacterial PCA. Two semi-random leucine zipper libraries were created (as described in the text) and each inserted N-terminal to one of the mDHFR fragments. Cotransformation of the resulting zipper-DHFR fragment libraries in *E. coli* and plating on selective medium allowed for survival of clones harboring successfully interacting leucine zippers. Fourteen clones were isolated and the zippers were sequenced to identify the residues at the OeO and OgO positions. The Oe-gO pairs were categorized, as having attractive pairing (charge:charge, charge:neutral polar or neutral polar:neutral polar) or repulsive pairing (charge:charge) and the number of each type of interaction scored for each clone. The total number of interactions for each clone is 6; the interactions are tallied on the histogram.

Other objects, advantages and features of the present invention will become more apparent upon reading of the following non-restrictive description of preferred embodiments with reference to the accompanying drawings which are exemplary and should not be interpreted as limiting the scope of the present invention.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

Selection of mDHFR for a PCA

In designing a protein-fragment complementation assay (PCA), we sought to identify an enzyme for which the following is true: 1) An enzyme that is relatively small and monomeric, 2) for which structural and functional information exists, 3) for which simple assays exist for both *in vivo* and *in vitro* measurement, and 4) for which overexpression in eukaryotic and prokaryotic cells has been demonstrated. Murine DHFR (mDHFR) meets all of the criteria for a PCA listed above. Prokaryotic and eukaryotic DHFR is central to cellular one-carbon metabolism and is absolutely required for cell survival in both prokaryotes and eukaryotes. Specifically it catalyses the reduction of dihydrofolate to tetrahydrofolate for use in transfer of one-carbon units required for biosynthesis of serine, methionine, purines and thymidylate. The DHFRs are small (17 kD to 21 kD), monomeric proteins. The crystal structures of DHFR from various bacterial and eukaryotic sources are known and substrate binding sites and active site residues have been determined¹¹¹⁻¹¹⁴, allowing for rational design of protein fragments. The folding, catalysis, and kinetics of a number of DHFRs have been studied extensively¹¹⁵⁻¹¹⁹. The enzyme activity can be monitored *in vitro* by a simple spectrophotometric assay¹²⁰, or *in vivo* by cell survival in cells grown in the absence of DHFR end products. DHFR is specifically inhibited by the anti-folate drug trimethoprim. As mammalian DHFR has a 12000-fold lower affinity for trimethoprim than does bacterial DHFR¹²¹, growth of bacteria expressing mDHFR in the presence of trimethoprim levels lethal to bacteria is an efficient means of selecting for reassembly of mDHFR fragments into

active enzyme. High level expression of mDHFR has been demonstrated in transformed prokaryote or transfected eukaryotic cells¹²²⁻¹²⁶.

Design Considerations.

5 mDHFR shares high sequence identity with the human DHFR (hDHFR) sequence (91% identity) and is highly homologous to the *E. coli* enzyme (29% identity, 68% homology) and these sequences share visually superimposable tertiary structure¹¹¹. Comparison of the crystal structures of mDHFR and hDHFR suggests that their active sites are essentially identical^{127,128}. DHFR has been described as being formed of 10 three structural fragments forming two domains^{129, 130} the adenine binding domain (residues 47 to 105 = fragment[2]) and a discontinuous domain (residues 1 to 46 = fragment[1] and 106 to 186 [3]; numbering according to the murine sequence). The folate binding pocket and the NADPH 15 binding groove are formed mainly by residues belonging to fragments[1] and [2]. Fragment [3] is not directly implicated in catalysis.

Residues 101 to 108 of hDHFR, at the junction between fragment[2] and fragment[3], form a disordered loop which lies on the same face of the protein as both termini. We chose to cleave mDHFR 20 between fragments [1,2] and [3], at residue 107, so as to cause minimal disruption of the active site and NADPH cofactor binding sites. The native *N*- terminus of mDHFR and the novel *N*-terminus created by cleavage occur on the same surface of the enzyme^{112,128} allowing for ease of *N*-terminal covalent attachment of each fragment to associating 25 fragments such as the leucine zippers used in this study. Using this system, we have obtained leucine-zipper assisted assembly of the mDHFR fragments into active enzyme.

EXAMPLE 1

EXPERIMENTAL PROTOCOL

DNA Constructs

5 Mutagenic and sequencing oligonucleotides were purchased from Gibco BRL. Restriction endonucleases and DNA modifying enzymes were from Pharmacia and New England Biolabs. The mDHFR fragments carrying their own iN-frame stop codon were subcloned into pQE-32 (Qiagen), downstream from and iN-frame with the
10 hexahistidine peptide and a GCN4 leucine zipper (Fig 1; Fig. 2). All final constructs were based on the Qiagen pQE series of vectors, which contain an inducible promoter-operator element (tac), a consensus ribosomal binding site, initiator codon and nucleotides coding for a hexahistidine peptide. Full-length mDHFR is expressed from pQE-16
15 (Qiagen).

Expression vector harboring the GCN4 leucine zipper

Residues 235 to 281 of the GCN4 leucine zipper (a Sall/BamHI 254 bp fragment) were obtained from a yeast expression
20 plasmid pRS316⁹. The recessed terminus at the BamHI site was filled-in with Klenow polymerase and the fragment was ligated to pQE-32 linearized with Sall/HindIII(filled-in). The product, construct Z, carries an open reading frame coding for the sequence Met-Arg-Gly-Ser followed by a hexahistidine tag and 13 residues preceding the GCN4 leucine zipper
25 residues.

Creation of DHFR fragments:

The eukaryotic transient expression vector, pMT3 (derived from pMT2)¹⁶, was used as a template for PCR-generation of mDHFR containing the features allowing subcloning and separate
5 expression of fragment[1,2] and fragment[3]. The megaprimer method of PCR mutagenesis²⁹ was used to generate a full-length 590 bp product. Oligonucleotides complementary to the nucleotide sequence coding for the N- and C-termini of mDHFR and containing a novel BspEI site outside the coding sequence were used as well as an oligonucleotide used to
10 create a novel stop codon after fragment[1,2], followed by a novel SpeI site for use in subcloning fragment[3].

Construction of a new multiple cloning region and subcloning of DHFR fragments [1,2] and [3]

15 Complementary oligonucleotides containing the novel restriction sites: SnaBI, NheI, SpeI and BspEI, were hybridized together resulting in 5' and 3' overhangs complementary to EcoRI, and inserted into pMT3 at a unique EcoRI site. The 590 bp PCR product (described above) was digested with BspEI and inserted into pMT3 linearized at
20 BspEI, yielding construct [1,2,3]. The 610 bp BspEI/EcoNI fragment (coding for DHFR fragment[1,2], followed by a novel stop and fragment[3] up to EcoNI) was filled in at EcoNI and subcloned into pMT3 opened with BspEI/HpaI, yielding construct F[1,2]. The 250 bp SpeI/BspEI fragment
25 of construct [1,2,3] coding for DHFR fragment[3] (with no in-frame stop codon) was subcloned into pMT3 opened with the same enzymes. The stop codon of the wild-type DHFR sequence, downstream from fragment[3] in pMT3, was inserted as follows. Cleavage with EcoNI,

present in both the inserted fragment[3] and the wild-type fragment[3], removal of the 683 bp intervening sequence and religation of the vector yielded a construct of fragment[3] with the wild-type stop codon, construct F[3].

5

Creation of the expression constructs

The 1051 bp and the 958 bp SnaBI/XbaI fragments of constructs F[1,2] and F[3], respectively, were subcloned into construct Z opened with BglII(filled-in)/NheI, yielding constructs Z-F[1,2] and Z-F[3] (Fig. 2). For the Control expression construct, the 180 bp XmaI/BspEI fragment coding for the zipper was removed from construct Z-F[1,2], yielding construct Control-F[1,2] (Fig. 2).

10

Creation of Stability Mutants

Site-directed mutagenesis was performed³⁰ to produce mutants at Ile114 (numbering of the wild-type mDHFR). The mutagenesis reaction was carried out on the KpnI/BamHI fragment of construct Z-F[3] subcloned into pBluescript SK+ (Stratagene), using oligonucleotides that encode a silent mutation producing a novel BamHI site. The 206 bp NheI/EcoNI fragment of putative mutants identified by restriction was subcloned back into Z-F[3]. The mutations were confirmed by DNA sequencing.

15

20

E. coli Survival Assay

25

E. coli strain BL21 carrying plasmid pRep4 (from Qiagen, for constitutive expression of the lac repressor) were made competent, transformed with the appropriate DNA constructs and washed

twice with minimal medium before plating on minimal medium plates containing 50 mg/ml kanamycin, 100 mg/ml ampicillin and 0.5 mg/ml trimethoprim. One half of each transformation mixture was plated in the absence, and the second half in the presence, of 1 mM IPTG. All plates
5 were placed at 37°C for 66 hrs.

E. coli Growth Curves

Colonies obtained from cotransformation were propagated and used to inoculate 10 ml of minimal medium supplemented with ampicillin, kanamycin as well as IPTG (1mM) and
10 trimethoprim (1 mg/ml) where indicated. Cotransformants of Z-F[1,2] + Z-F[3:Ile114Gly] were obtained under non-selective conditions by plating the transformation mixture on L-agar (+ kanamycin and ampicillin) and screening for the presence of the two constructs by restriction analysis.
15 All growth curves were performed in triplicate. Aliquots were withdrawn periodically for measurement of optical density. Doubling time was calculated for early logarithmic growth (OD 600 between 0.02 and 0.2).

Protein Overexpression and Purification

20 Bacteria were propagated in Terrific Broth³¹ in the presence of the appropriate antibiotics to an OD600 of approximately 1.0. Expression was induced by addition of 1 mM IPTG and further incubation for 3 hrs. For analysis of crude extract, pellets from 150 ml of induced cells were lysed by boiling in loading dye. The lysates were clarified by
25 microcentrifugation and analyzed by SDS-PAGE³². For protein purification, a cell pellet from 50 ml of induced *E. coli* cotransformed with constructs Z-F[1,2] and Z-F[3] was lysed by sonication, and a denaturing

purification of the insoluble pellet undertaken using Ni-NTA (Qiagen) as described by the manufacturer. The proteins were eluted with a stepwise imidazole gradient. The fractions were analyzed by SDS-PAGE.

RESULTS

Design of mDHFR fragments for a PCA

mDHFR shares high sequence identity with the human DHFR (hDHFR) sequence. As the coordinates of the murine crystal structure were not available, we based our design considerations on the hDHFR structure. DHFR has been described as comprising three structural fragments forming two domains: the adenine binding domain (F[2]) and a discontinuous domain (F[1] and F[3])^{13,18}. The folate binding pocket and the NADPH binding groove are formed mainly by residues belonging to F[1] and F[2]. Residues 101 to 108 of hDHFR form a disordered loop which lies on the same face of the protein as both termini. This loop occurs at the junction between F[2] and F[3]. By cleaving mDHFR at residue 107, we created F[1,2] and F[3], thus causing minimal disruption of the active site and substrate binding sites. The native *N*-terminus of mDHFR and the novel *N*-terminus created by cleavage were covalently attached to the *C*-termini of GCN4 leucine zippers (Fig. 1).

E. coli Survival Assays

Figure 2 illustrates the general features of the expressed constructs and the nomenclature used in this study. Figure 3 (panel A) illustrates the results of cotransformation of bacteria with constructs coding for Z-F[1,2] and Z-F[3], in the presence of trimethoprim, clearly showing that colony growth under selective pressure is possible only in

cells expressing both fragments of mDHFR. There is no growth in the presence of either Z-F[1,2] or Z-F[3] alone. Induction of protein expression with IPTG is essential for colony growth (Fig. 3A). The presence of the leucine zipper on both fragments of mDHFR is essential as illustrated by cotransformation of bacteria with both vectors coding for mDHFR fragments, only one of which carries a leucine zipper (Fig. 3A). It should be noted that growth of control *E. coli* transformed with the full-length mDHFR is possible in the absence of IPTG due to low levels of expression in uninduced cells.

Confirmation of the presence of both plasmids in bacteria able to grow with trimethoprim was obtained from restriction analysis of the plasmid DNA purified from isolated colonies. Figure 4 (A) reveals the presence of the 1200 bp HincII restriction fragment from construct Z-F[1,2] as well as the 487 and 599 bp HincII restriction fragments from construct Z-F[3]. Also present is the 935 bp HincII fragment of pRep4. Overexpression of the fusion proteins is illustrated in Figure 4 (B). In all cases, overexpression of a protein of the expected molecular weight is apparent on SDS-PAGE of the crude lysate. Purification of the coexpressed proteins under denaturing conditions yielded two bands of apparent homogeneity upon analysis by Coomassie-stained SDS-PAGE (Fig. 4B).

Stability Mutants

Applicants generated mutants of F[3] to test whether reconstitution of mDHFR activity by fragment assembly was specific. Protein stability can be reduced by changing the side-chain volume in the hydrophobic core of a protein^{9, 22-25}. Residue Ile114 of mDHFR occurs in

a core b-strand at the interface between F[1,2] and F[3], isolated from the active site. Ile 114 is in van der Waals contact with Ile51 and Leu93 in F[1,2]¹¹. We mutated Ile 114 to Val, Ala, or Gly. Figure 3 (panel B) illustrates the results of cotransformation of *E. coli* with construct Z-F[1,2] and the mutated Z-F[3] constructs. The colonies obtained from cotransformation with Z-F[3:Ile114Ala] grew more slowly than those cotransformed with Z-F[3] or Z-F[3:Ile114Val] (see inset to Fig. 3B). No colony growth was detected in cells cotransformed with Z-F[3:Ile114Gly]. The number of transformants obtained was not significantly different in the case where colonies were observed, implying that cells cotransformed with Z-F[1,2] and either Z-F[3], Z-F[3:Ile114Val] or Z-F[3:Ile114Ala] have an equal survival rate. Overexpression of the mutants Z-F[3:Ile114X] was in the same range as Z-F[3], as determined by Coomassie-stained SDS-PAGE (data not shown).

We also compared the relative efficiency of reassembly of mDHFR fragments by measuring the doubling time of the cotransformants in liquid medium. Doubling time in minimal medium was constant for all transformants (data not shown). Selective pressure by trimethoprim in the absence of IPTG prevented growth of *E. coli* except when transformed with pQE-16 coding for full-length DHFR due to low levels of expression in uninduced cells. Induction of mDHFR fragment expression with IPTG allowed survival of cotransformed cells (except in the case of Z-F[1,2] + Z-F[3:Ile114Gly], although the doubling times were significantly increased relative to growth in the absence of trimethoprim. The doubling time measured for cells expressing Z-F[1,2] + Z-F[3], Z-F[1,2] + Z-F[3:Ile114Val] and Z-F[1,2] + Z-F[3:Ile114Ala] were 1.6-fold, 1.9-fold and 4.1-fold, higher respectively, than the doubling time of *E. coli*

expressing pQE-16 in the absence of trimethoprim and IPTG. The presence of IPTG unexpectedly prevented growth of *E. coli* transformed with full-length mDHFR. Growth was partially restored by addition of the folate metabolism end-products thymine, adenine, pantothenate, glycine and methionine (data not shown). This suggests that induced overexpression of mDHFR was lethal to *E. coli* when grown in minimal medium as a result of depletion of the folate pool by binding to the enzyme.

In another embodiment, applicants make point mutations in the GCN4 leucine zipper of Z-F[1,2] and Z-F[3], for which direct equilibrium and kinetic parameters are known and correlating these known values with parameters derived from the PCA (Pelletier and Michnick, in preparation). Comparison of cell growth rates in this model system with rates for a DHFR PCA using unknowns would give an estimate of the strength of the unknown interaction. This should enable the determination of estimates of equilibrium and kinetic parameters for a specific protein-protein interaction.

The present invention has illustrated and demonstrated a protein-fragment complementation assay (PCA) based on mDHFR, where a leucine zipper directs the reconstitution of DHFR activity. Activity was detected by an *E. coli* survival assay which is both practical and inexpensive. This system illustrates the use of mDHFR fragment complementation in the detection of leucine zipper dimerization and could be applied to the detection of unknown, specific protein-protein interactions *in vivo*.

***E. coli* Aminoglycoside kinase: Optimization and Design of a PCA using an Exonuclease-Molecular Evolution Strategy**

Although applicants have demonstrated that the engineering/design strategy described above can be used to produce complementary enzyme fragments, it is obvious that proteins did not evolve in such a way that such fragments would be expected to have optimal physical characteristics, including solubility, foldability (fast folding), protease resistance, or enzymatic activity. An alternative embodiment to the engineering/design strategy is the endonuclease/evolution approach. This strategy can be used by itself or in conjunction with the engineering/design strategy. The advantages of this approach are that in principle, prior knowledge of the protein structure is not necessary, that the optimal fragments are chosen for PCA and that these fragments will also have optimal characteristics. Following selection of optimal complementary fragments, the fragments are exposed to multiple rounds of random mutagenesis. Mutagenesis is achieved by suboptimal PCR combined with chemical mutagenesis or DNA shuffling (Stemmer, W. P. C. (1994) Proc, Natl, Acad, Sci. USA 91, 10747-10751). The overall strategy is described for the case of aminoglycoside kinase (AK), an example of antibiotic resistance marker that can be used for dominant selection of prokaryotic cells such as *E. coli* or eukaryotic cells such as yeast or mammalian cell lines. The structure of an AK is already known, and so strategy (1) would be possible, however we chose to combine both strategy (1) as defined for DHFR above, in conjunction with strategy (2).

EXPERIMENTAL PROTOCOL

The optimization/selection procedure is as follows:

Generation of of library of AK fragments based on products of

5 Exonuclease digestion

Nested sets of deletions are created at the 5' and the 3' ends of the AK gene. In order to create unidirectional deletions, unique restriction sites are introduced in the regions flanking the AK gene. At the 5' and 3' termini, an "outer" sticky site with a protruding 3' terminus (Sph I and Kpn I, respectively) and an "inner" sticky site with recessed 3' terminus (Bgl II and Sal I, respectively) are added by PCR. Cleavage at Sph I and Bgl II (or Kpn I and Sal I) results in creation of a protruding terminus leading back to the flanking sequence and a recessed terminus leading into the AK gene. Digestion with *E. coli* exonuclease III and S1 nuclease (Henikoff, S. (1987) Methods in Enzymology 155, 156-165) yields a set of nested deletions from the recessed terminus only. Thus, 10 mg of DNA is digested with Sph I and Bgl II (or Kpn I and Sal I), phenol-chloroform extracted, and 12.5 U exonuclease III added. At 30 sec intervals over 10 min, aliquots are taken and put into solution with 2 U S1 nuclease. The newly created ends are filled in with T4 DNA polymerase (0.1 U per sample) and the set of vectors closed back by blunt-ended ligation (10 U ligase per sample). The average length of the deletion at each time point is determined by restriction analysis of the sets. This yields sets of AK genes deleted from the 5' or the 3' termini. This manipulation is undertaken directly in the pQE-32-Zipper constructs, such that the products can be used directly in activity screening.

Screening for AK activity

As a first step in determining the requirements for fragment complementation, we must determine the minimum *N*-terminal and *C*-terminal fragments of AK that, alone, are active. Sets of deletions are individually transformed into *E. coli* BL21 cells and expression of the AK fragments is induced by IPTG. The sets where a significant number of colonies appear in the presence of G418 serve to indicate the approximate length of *N*- and *C*-terminal AK fragments which retain activity. Fragment complementation must therefore be undertaken with fragments taken from within these limits. The zipper-directed fragment complementation is detected as follows: appropriate sets of deletions, or pools of sets, are cotransformed into BL21, expression is induced with IPTG and growth in the presence of varying G418 concentrations is monitored. Large colonies which grow in the presence of high G418 concentrations are selected as giving the most efficiently complementing products.

Directed evolution of optimal AK fragments using "DNA shuffling"

After optimal fragments have been selected, the individual fragments are removed by restriction digestion at *Sph* I and *Kpn* I allowing for 5' and 3' constant priming regions flanking the *N*- or *C*-terminal complementary fragments of AK. These oligonucleotides (2-4 mg) are digested with *DNase*I (0.005 units/ μ l, 100 μ l) and fragments of 10-50 nucleotides are extracted from low melting point agarose. PCR is then performed with the fragmented DNA, using *Taq* polymerase (2.5 units/ μ l) in a PCR mixture containing 0.2 mM dNTPs, 2.2 mM $MgCl_2$ (or 0 mM for suboptimal PCR), 50 mM KCl, 10 mM Tris.HCl, pH 9.0, 0.1%

TritonX-100. A PCR program of 94C/60 sec.; 94C 30 sec.; 55C 30 sec.; 72C 30 sec. times 30 to 50; 72C 5 min. Samples are taken every 5 cycles after 25 cycles to monitor the appearance of reassembled complete fragments on agarose gel. The primerless PCR product is then
5 diluted 1:40 or 1:60 and used as template for PCR with 5', 3' complementary constant region oligos as primers for a further 20 cycles. Final product is restriction digested with Sph I and Kpn I and the products subcloned back into pQE32-Zipper to yield the final library of expression plasmids. As before, *E. coli* BL21 cells are sequentially transformed with
10 C-terminal or N-terminal complementary fragment-expression vectors at an estimated efficiency of 10⁹ and finally cells cotransformed with the complementary fragment. *E. coli* are grown on agarose plates containing 1 mg/ ml G418 and after 16 hours the largest colonies are selected and grown in liquid medium at increasing concentrations of G418. Those
15 clones showing the maximal resistance to G418 are then selected and if maximum resistance or greater is reached the evolution is terminated. Otherwise the DNA shuffling procedure is repeated. Finally, optimal fragments are sequenced and physical properties and enzymatic activity are assessed. This optimized AK PCA is now ready to test for dominant
20 selection in any other cell type including yeast and mammalian cell lines. This strategy can be used to develop any PCA based on enzymes that impart dominant or recessive selection to a drug or toxin or to enzymes that produce a colored or fluorescent product. In the later two cases the end point of the evolution process is at minimum, reattainment of signal
25 for the intact, wild type enzyme or enhancement of the signal. This strategy can also be used in the absence of knowledge of the enzyme structure, whether the enzyme in mono-, di- or multimeric structure.

However, knowledge of the enzyme structure does not preclude applying this strategy as well, as described below.

As can be appreciated, knowledge of the enzyme structure can be used to render a more efficient way of using molecular evolution to design a PCA. In this case, the enzyme structure is used to define minimal domains of the protein in question, as was done for DHFR. Instead of generating fragments of completely random length for the *N*- or *C*-terminal fragments, we select, during the exonuclease phase, those fragments that at a minimum will code for one of the two domains. For instance, in the case of AK, two well defined domains can be discerned in the structure consisting of residues 1-94 in the *N*-terminus and residues 95-267 in the *C*-terminus. Endonuclease digestions are performed as above, but reaction products are selected that will minimally code for one of the two domains. These are then the starting points for fragment selection and evolution cycles as described above.

Heteromeric Enzyme PCA

A further embodiment of the invention relates to PCA based on using heterodimeric or heteromultimeric enzymes in which the entire catalytic machinery is contained within one independently folding subunit and the other subunit provides stability and/or a cofactor to the enzymatic subunit. In this embodiment of PCA, the regulatory subunit is split into complementary fragments and fused to interacting proteins. These fragments are co-transformed/transfected into cells along with the enzyme subunit. As with single enzyme PCA described for DHFR and AK, reconstitution and detection of enzyme activity is dependent on oligomerization domain-assisted reassembly of the regulatory subunit

reassembly into its native topology. However, the reconstituted subunit then interacts with the intact enzymatic subunit to produce activity. This approach is reminiscent of the USPS system, except it has the advantage that the enzyme in this case is not a constitutive cellular enzyme, but rather an exogenous gene product. As such there is no problem with background activity from the host cell, the enzyme can be expressed at higher levels than a natural gene and can also be modified to be directed to specific subcellular compartments (by subcloning compartment-specific signal peptides onto the *N*- or *C*-termini of the enzyme and subunit fragments). The specific advantage of this approach is that while the single enzyme strategy may lead to suboptimal enzymatic activity, in this approach, the enzyme folds independently and may in fact act as a chaperone to the fragmented regulatory subunit, aiding in its refolding. In addition, folding of the fragments may need not be complete in order to impart regulation of the enzyme. This approach is realized by a colorimetric/fluorometric assay we have developed based on the *Streptomyces* tyrosinase. This enzyme catalyzes the conversion of tyrosine to deoxyphenylalanine (DOPA). The reaction can be measured by conversion of fluorocinyl-tyrosine to the DOPA form. The active enzyme consists of two subunits, the catalytic domain (Melc2) and a copper binding domain (Melc1). Melc1 is a small protein of 14 kD that is absolutely required for Melc2 activity. In the assay we are developing, the Melc1 protein is split into two fragments that serve as the complementation part of the PCA. These fragments, fused to oligomerization domains, are coexpressed with Melc2, and the basis of the assay is that Melc2 activity is dependent on complementation of the Melc1 fragments. Stoichiometries of protein complexes can also be

addressed (i.e. whether a complex consists of two or three proteins) as follows. One fuses two proteins to the two Melc1 fragments and a third to intact Melc2. It thus can be shown that the minimum complementary active complex of the tyrosinase will require that all three components and therefore a trimer is necessary. A key aspect of this approach is that we can easily demonstrate specific interactions by making one component, specifically the protein-Melc2 fusions catalytic subunit dependent on the other components by underexpressing it in the background of overexpressed Melc1 fragment-protein fusions.

10

Multimer Disruption-Based PCA

Although applicants have described only fragment complementation of intact proteins, protein domains or subunits as comprising PCA, an alternate embodiment relates to PCAs based on the disruption of the interface between, for instance a dimeric enzyme that requires stable association of the subunits for catalytic activity. In such cases, selective or random mutagenesis at the subunit interface would disrupt the interaction and the basis of the assay would be that oligomerization domains fused to the subunits would provide the necessary binding energy to bring the subunits together into a functional enzyme.

20

Vector Design in Application to PCAs

The PCA strategies listed thus far have used two-plasmid transformation strategies for expression of complementary fragments. This approach has some advantages, such as using different drug resistance markers to select for optimal incorporation of genes, for instance in transformed or transfected cells or for optimum transformation

25

of complementary plasmids into bacteria and control of expression levels of PCA fragments using different promoters. However, single plasmid strategies have advantages in terms of simplicity of transfection/transformation. Protein expression levels can be controlled in different ways, while drug selection can be achieved in one of two ways: In the case of PCAs based on survival assay using enzymes that are drug resistance markers themselves, such as AK, or where the enzyme complements a metabolic pathway, such as DHFR, no additional drug resistance genes need be incorporated in the expression plasmids. If however the PCA is based on an enzyme that produces a colored or fluorescent product, such as tyrosinase or firefly luciferase, an additional drug resistance gene must be expressed from the plasmid. Expression of PCA complementary fragments and fused cDNA libraries/target genes can be assembled on single plasmids as individual operons under the control of separate inducible or constitutive promoters, or can be expressed polycistronically. In *E. coli* polycistronic expression can be achieved using known intercoding region sequences, for instance we use the region in the *mel* operon from which we derived the tyrosinase *melc1-melc2* genes which we have shown to be expressed at high levels in *E. coli* under the control of a strong (*tac*) promoter. Genes could also be expressed and induced off of independent promoters, such as *tac* and arabinose. For mammalian expression systems, single plasmid systems can be used for both transient or stable cell line expression and for constitutive or inducible expression. Further, differential control of the expression of one of the complementary fragment fusions, usually the bait-fused fragment, can be controlled to minimize expression. This will be important in reducing background non-specific interactions. Examples

of differential control of complementary fragment expression include the following strategies:

- 5 i) In polycistronic expression, transient or stable, expression of the second gene will necessarily be less efficient and so this in itself could serve to limit the quantity of one of the complementary fragments. Alternatively, the first gene product can be limited in expression by mutation of an upstream donor/splice site, while the second gene can be put under the control of a retroviral internal initiation site, such as that of ECMV to enhance expression.
- 10 ii) Individual complementary fragment-fusion pairs can also be put under the control of inducible promoters, all commercially available including those based on Tet-responsive PhCMV*-1 promoter, and/or steroid receptor response elements. In such a system the two complementary fragment genes can be turned on and expression levels controlled by
- 15 dose dependent expression with the inducer, in these cases tetracycline and steroid hormones.

EXAMPLE 2

20 Applications of the PCA strategy to detect novel gene products in biochemical Pathways and to map such pathways

Among the greatest advantage of PCA over other molecular interaction screening methods is that they are designed to be performed both *in vivo* and in any type of cell. This feature is crucial if the

25 goal of applying a technique is to identify novel interactions from libraries and simultaneously be able to determine if the interactions observed are biologically relevant. The detailed example given below, and other

examples at the end of this section illustrate how it is that validation of interactions with PCA is possible. In essence, this is achieved as follows. In biochemical pathways, such as hormone receptor-mediated signaling, a cascade of enzyme-mediated chemical reactions are triggered by some

5 molecular event, such as by hormone binding to its membrane surface receptor. Enzyme interactions with protein substrates and protein-protein or protein-nucleic acid interactions with enzyme-modified substrates then occur. Such biochemical signaling cascades only occur in specific cell types and model cell lines for studying these processes. Therefore, to

10 detect induced interactions, such as with known proteins in a pathway with yet unidentified proteins, one obviously needs to perform such screening in appropriate model cell lines and in the correct cellular compartment. Only the PCA strategy can be used in a general way to do this. Protein-molecular interaction techniques such as yeast two- or

15 three-hybrid techniques cannot be performed in a context where such events occur, except in the limiting case of nuclear interaction in yeast or interactions that are not triggered. There do exist mammalian two-hybrid techniques where it might be possible to detect induced protein interactions, but only again if the proteins involved can be simultaneously

20 activated, transported to the nucleus and interact with their partners. PCAs do not have these limitation since they do not require additional cellular machinery available only in specific compartments. A further point is that by performing the PCA strategy in appropriate model cell types, it is also possible to introduce appropriate positive and negative

25 controls for studying a particular pathway. For instance, for a hormone signaling pathway it is likely that hormone signaling agonists and antagonists or dominant-negative mutants of signaling cascade proteins

would be known, that are upstream or act in parallel to the events being examined in the PCA. These reagents could be used to determine if novel interactions detected by the PCA are biologically relevant. In general then, interactions that are detected only if hormone is introduced but are not seen if an antagonist is simultaneously introduced could be hypothesized to represent interactions relevant to the process under study.

Below is a detailed description of an application of the DHFR that illustrates these points, as well as further examples where the PCA strategy could be used.

Application of the DHFR PCA to Mapping Growth Factor-Mediated Signal Transduction Pathways

One of the earliest detectable events in growth factor-activated cell proliferation is the serine phosphorylation of the S6 protein of the 40S ribosomal subunit. The discovery of serine/threonine kinases that specifically phosphorylate S6 have considerably aided in identifying novel mitogen mediated signal transduction pathways. The serine/threonine kinase p70S6k has been identified as a specific S6 phosphorylase¹³¹⁻¹³⁶. p70S6k is activated by serine and threonine phosphorylation at specific sites in response to several mitogenic signals including serum in serum starved cells, growth factors including insulin and PDGF, and by mitogens such as phorbol esters. Considerable effort has been made over the last five years to determine how p70/p85S6k are activated in response to mitogens. Two receptor-mediated pathways have been implicated in p70S6k activation, one associated with the phosphatidylinositol-3-kinase (PI(3)k) and the other with the PI(3)k

homologue mTOR¹³⁷⁻¹⁴⁴. Key to understanding of this proposal, is the fact that the role of these enzymes in activation of p70S6k was determined by effects of two natural products on phosphorylation and enzyme activity: rapamycin, which indirectly inhibits mTOR activity, and wortmannin, which
5 directly inhibits PI(3)k activity. It is also important to note that no direct upstream kinases or other regulatory proteins of p70S6k have been identified to this date.

The interactions of p70S6k with its known substrate S6 can be studied as a test system for the DHFR PCA in *E. coli* and in
10 mammalian cell lines. One can also seek to identify novel interactions with this enzyme that would lead to new insights into how this important enzyme is regulated. Also, since activation of the enzyme is mediated by multiple pathways that can be selectively inhibited with specific drugs, this is an ideal system to test PCAs as methods to distinguish induced versus
15 constitutive protein-protein interactions.

a) Testing of the *E. coli* survival assay: Interaction of p70S6k with S6

This test is ideal, because the apparent K_m (= 250 nM) of p70S6k for S6 protein¹⁴⁵ is approximately the same as the K_d for leucine zipper-forming peptides from GCN4 used in our test system.
20 However, we will have to use a constitutively active form of the enzyme for our tests. An N-terminal truncated form of the enzyme D77-p70S6k, is constitutively active and will be used in these studies¹⁴⁷.

Methodology: D77-p70S6k-F[1,2] fusion and D77-p70S6k-F[3] fusion, or
25 F[1,2] and D77-p70S6k-F[3] fusion (as a control) will be cotransformed into *E. coli* and the cells grown in minimal medium in the presence of trimethoprim. Colonies will be selected and expanded for analysis of

kinase activity against 40S ribosomal subunits, and for coexpression of the two proteins.

b) Modification of the bacterial survival assay for library screening:

5 **Identification of Novel Interacting Proteins**

Screening an expression library for interactions with a given target (p70S6k-D77, in this case) will be straightforward in this system, given that the only steps involved are: 1-construction of the fusion-expression library as a fusion with mDHFR fragment[3]; 2-
10 transformation of the library in *E. coli* BL21 harboring pRep4 (for constitutive expression of the lac repressor; this is required in the case where a protein product is toxic to the cells) and a plasmid coding for the fusion: p70S6k-D77-[1,2]; 3-plating on minimal medium in the presence of trimethoprim and IPTG; 4-selection of any colonies that grow,
15 propagation and isolation of plasmid DNA, followed by sequencing of DNA inserts; 5-purification of unknown fusion products via the hexaHis-tag and sizing on SDS-PAGE.

Methodology:

20 The overall strategy is illustrated in Figure 5. 1- Construction of a directional fusion-expression library: i-cDNA production: One can isolate poly(A)+ RNA from BA/F3 cells (B-lymphoid cells) because these cells have successfully been used in the study of the rapamycin-sensitive p70S6k activation cascade¹³⁹. To enrich for full-
25 length mRNA, we will affinity purify the mRNA via the 5' cap structure by the CAPture method¹⁴⁸. Reverse transcription will be primed by a "Linker Primer": it has a poly(T) tail to prime from the poly(A) mRNA tail, and an

XhoI site for later use in directional subcloning of the fragments. The first strand is then methylated. After second strand synthesis and blunting of the products, "EcoRI Adapters" are added, producing digestion of the linkers with EcoRI and XhoI (the inserts are protected by methylation) produces full-length cDNA ready for directional insertion in a vector opened with EcoRI and XhoI. Because the success of library screening depends largely on the quality of the cDNA produced, we will use the above methods as they have proven to consistently produce high-quality cDNA libraries. ii-Insertion of the cDNA into vectors: The library will be constructed as a C-terminal fusion to mDHFR F[3] in vector pQE-32 (Qiagen), as we have obtained high levels of expression of mDHFR fusions from this vector in BL21 cells. Three such vectors will be created, differing at their 3' end, which is the novel polycloning site that we engineered (described earlier, under Methods), carrying either 0, 1, or 2 additional nucleotides. This allows read-through from F[3] into the library fragments in all 3 translational reading frames. The cDNA fragments will be directionally inserted at the EcoRI and XhoI sites in all three vectors at once. 2, 3, 4, and 5- These steps have been described earlier, under Results, apart from the final sequencing of clones identified using sequencing primers specific to vector sequences flanking sites of library insertion. The protein purification will also be as described earlier, by a one-step purification on Ni-NTA (Qiagen). If the product size is more than 15 kDa over the molecular weight of the DHFR component (equal to a cDNA insert of more than 450bp), we will have the inserts sequenced at the Sheldon Biotechnology Center (McGill University).

c) Development of the Eukaryotic Assay

The transformation of the system described above, is useful to produce an equivalent assay for use in eukaryotic cells. The basic principle of the assay is the same: the fragments of mDHFR are fused to associating domains, and domain association is detected by reconstitution of DHFR activity in eukaryotic cells (Figure 5).

Creation of the expression constructs: The DNA fragments coding for the GCN4-zipper-mDHFR fragment fusions were inserted as one piece into pMT3, a eukaryotic transient expression vector¹²⁶. Expression of the fusion proteins in COS cells was apparent on SDS-PAGE after 35[S]Met labeling.

Survival assays in eukaryotic cells: Two systems can be used for detection of mDHFR reassembly, in parallel: i- CHO-DUKX B11 cells (Chinese Hamster Ovary cell line deficient in DHFR activity) are cotransfected with GCN4-zipper-mDHFR fragment fusions. The cells are grown in the absence of nucleotides; only cells carrying reconstituted DHFR will undergo normal cell division and colony formation. ii- Methotrexate (MTX)-resistant mutants of mDHFR have been created, with the goal of transfecting cells that have constitutive DHFR activity such as COS and 293 cells. We mutated F[1,2] in order to incorporate, one at a time, each of five mutations that significantly increase K_i (MTX): Gly15Trp, Leu22Phe, Leu22Arg, Phe31Ser and Phe34Ser (numbering according to the wild-type mDHFR sequence). These mutations occur at varying positions relative to the active site and relative to F[3], and have varying effects on K_m (DHF), K_m (NADPH) and V_{max} of the full-length mammalian enzymes in which they were. Mutants Z-F[1,2: Leu22Phe], Z-F[1,2: Leu22Arg] and Z-F[1,2: Phe31Ser] all allowed for bacterial

survival with high growth rates when cotransformed with Z-F[3] (results not shown). The five mutants will be tested in eukaryotic cells, in reconstitution of mDHFR fragments to produce enzyme that can sustain COS or 293 cell growth while under the selective pressure of MTX, which will eliminate background due to activity of the native enzyme. The mutations offers an advantage in selection while presenting no apparent disadvantage with respect to reassembly of active enzyme. If the reconstituted mDHFR produced in either of the survival assays allows eukaryotic cell growth that is significantly slower than growth with the wild-type enzyme, thymidylate will be added to the growth medium to partially relieve the selective pressure offered by the lack of nucleotides.

d) Testing of the eukaryotic survival assay

It is necessary at the outset to test whether induced interactions with p70S6k can be detected. One can use the same test system as that for the *E. coli* test system described above: Induction of association of p70S6k with S6 protein.

Methodology:

mDHFR Leu22Phe mutant S6-F[1,2] and p70S6k-F[3], or F[1,2] and p70S6k-F[3] (as a control) will be cotransfected into COS cells and the cells will be serum starved for 48 hours followed by replating of cells at low density in serum and MTX. Colonies will be selected and expanded for analysis of kinase activity against 40S ribosomal subunits, and for coexpression of the two proteins. Further controls will be performed for inhibition of protein association with wortmannin and rapamycin.

e) Modification of the eukaryotic survival assay for library screening

An important part of the work required in creating a library for use in eukaryotic cells will have been accomplished already, as

the EcoRI/XhoI directional cDNA produced by the Stratagene "cDNA Synthesis Kit" can directly be inserted directionally into the Stratagene Zap Express vector.

Methodology:

5 Steps 1 through 5 are parallel to those for the bacterial library screening (above). 1-Again, the library is constructed as a C-terminal fusion to mDHFR F[3]. F[3] (with no stop codon) will be inserted in frame in Zap Express, followed by insertion of the novel polylinkers allowing expression of the inserts in all three reading frames (described
10 above), and by the EcoRI/XhoI directional cDNA. This bacteriophage library will be propagated and treated with the Stratagene helper phage to excise a eukaryotic expression phagemid vector (pBK-CMV) carrying the fusion inserts. 2-Cotransfection of the library and p70S6k-F[1,2] constructs in eukaryotic cells: we will perform the screening in COS or
15 293 cells, as these are responsive to serum in activating the p70S6k signaling pathway. Selection experiments will be performed as described for the S6 test system above. 3-Propagation, isolation and sequencing of the insert DNA will be undertaken. 4-The cloned fusion proteins will be
20 sized on SDS-PAGE by direct visualization after 35S-Met/Cys labeling, or by Western blotting using a commercial polyclonal antibody to mDHFR.

Generalization of the Strategy: The scheme for detecting partners for the protein p70S6k can be applied to studies of any biochemical pathway in any living organism. Such pathways may also be related to disease processes. The disease-related pathway may be an intrinsic process of
25 cells in humans where a pathology arises from, for instance mutation, deletion or under or over expression of a gene. Alternatively the biochemical pathway may be one that is specific to a pathogenic

organism or the mechanism of host invasion. In this case, component proteins of such processes may be targets of a therapeutic strategy, such as development of drugs that inhibit invasion by the organism or a component enzyme in a biochemical pathway specific to the pathogenic organism.

Inflammatory diseases are a case in point that can concern both examples. The protein-protein interactions that mediate the adhesion of leukocytes to inflamed tissues are known to involve such proteins as vascular cell adhesion molecule-1 (VCAM-1), and certain cytokines such as IL-6 and IL-8 that are produced during inflammation. However, many of the proteins involved in onset of inflammatory response remain unknown; further, the intracellular signaling pathways triggered by the extracellular associations are poorly understood. The PCAs could be used in elucidation of the mechanisms underlying the onset of inflammation, as well the ensuing signaling. For example, signaling pathways associated with inflammation, such as those mediated by IL-1, IL-6, IL-8 and tumor necrosis have been studied in some detail and many direct and downstream regulators are known. These regulators can be used as starting point targets in a PCA screening to identify other signalling or modulating proteins that could also be targets for drug development.

There is an increased risk of infection by enteric pathogens in the occurrence of the intestinal inflammation that characterizes idiopathic intestinal diseases. There are two mechanisms which need to be better understood here and which can be addressed by PCA:

i- the cellular mechanisms of inflammation as described above, and

ii- the discovery of the specific cell-surface ligands which the pathogenic organisms recognize and associate with. Secreted proteins produced by the pathogen can bind to the basolateral membrane of epithelial cells (as in the case in *Yersinia pseudotuberculosis* infection) or be translocated
5 into intestinal epithelial cells (*Salmonella* infection), promoting infectivity and/or physiological responses to the infection. However, in most cases the interactions between the pathogenic protein and the epithelial cells are unknown.

Cell adhesion and nervous system regeneration A related example in
10 cell adhesion includes processes involved in development and regeneration in the nervous system. Cadherens are membrane proteins that mediate calcium dependent cell-cell adhesion. To do so they need another class of cytoplasmic proteins called catenins. Those make a bridge between cadherins and cytoskeleton. Catenins also regulate
15 genes that control differentiation-specific genes. For instance, the protein B-catenin can interact in certain situation with a transcription factor (lef-1) and be translocated into the nucleus where it constrains the number of genes transactivated by lef-1 (differentiation). This process is regulated by the Wnt signaling pathway (homologs to the wingless pathway in
20 *Drosophila*) by inactivation of GSK3B which permit degradation after of APC (a cytoplasmic adapter protein). PCA strategies could be used to identify novel proteins involved in the regulation of these processes.

Proteins involved in viral integration processes are examples of targets that could be tested for inhibitors using the PCA
25 strategies. Examples for the HIV virus include:

- i) inhibition of integrase or the transport of the pre-integration complex: protein Ma or vpr.
- ii) Inhibition of the cell cycle in G2 by vpr (interaction by cyclin B) causing induction of apoptosis.
- 5 iii) Inhibition of the interaction of gp160 (precursor of the membrane proteins) with furine.

Accessory proteins of HIV as a therapeutic target:

- 10 i) Vpr: nuclear localizing sequence (target): interaction site of vpr with phosphatasesA .
- ii) vif: interaction with vimentin (cytoskeleton associated protein).
- ii) Vpu: Degradation of CD4 in the RE mediated by the cytoplasmic tail of Vpu.
- iii) nef: Myristoylation signal of Nef.

15

EXAMPLE 3

Other Examples of Protein Fragment Complementation Assays

Other examples of assays are herein exemplified. The reason to produce these assays is to provide alternative PCA strategies that would be appropriate for specific protein association problems such as studying equilibrium or kinetic aspects of assembly. Also, it is possible that in certain contexts (for example, specific cell types) or for certain applications, a specific PCA will not work but an alternative one will. Further below are brief descriptions of each other PCAs embodiments.

20

1) **Glutathione-S-Transferase (GST)** GST from the flat worm *Schistosoma japonicum* is a small (28 kD), monomeric, soluble protein that can be expressed in both prokaryotic and eukaryotic cells. A high resolution crystal structure has been solved and serves as a starting point for design of a PCA. A simple and inexpensive colorimetric assay for GST activity has been developed consisting of the reductive conjugation of reduced glutathione with 1-chloro-2,4-dinitrobenzine (CNDNB), a brilliant yellow product. We have designed a PCA based on similar structural criteria used to develop the DHFR PCA using GCN4 leucine zippers as oligomerization domains. Cotransformants of zipper-GST-fragment fusions are expressed in *E. coli* on agar plates and colonies are transferred to nitrocellulose paper. Detection of fragment complementation is detected in an assay where a glutathione-CDNB reaction mixture is applied as an aerosol on the nitrocellulose and colonies expressing co-expressed fragments of GST are detected as yellow images.

2) **Green Fluorescent Protein (GFP)** GFP from *Aequorea victoria* is becoming one of the most popular protein markers for gene expression. This is because the small, monomeric 238 amino-acids protein is intrinsically fluorescent due to the presence of an internal chromophore that results from the autocatalytic cyclization of the polypeptide backbone between residues Ser65 and Gly67 and oxidation of the γ -bond of Tyr66. The GFP chromophore absorbs light optimally at 395 nm and possesses also a second absorption maximum at 470 nm. This bi-specific absorption suggests the existence of two low energy conformers of the chromophore whose relative population depends on local environment of the

chromophore. A mutant Ser65Thr that eliminates isomerization (single absorption maximum at 488 nm) results in a 4 to 6 times more intense fluorescence than the wild type. Recently the structure of GFP has been solved by two groups, making it now a candidate for a structure-based PCA-design, which we have begun to develop. As with the GST assay, we are doing all of our initial development in *E. coli* with GCN4 leucine zipper-forming sequences as oligomerization domains. Direct detection of fluorescence by visual observation under broad spectrum UV light will be used. We will also test this system in COS cells, selecting for co-transfectants using fluorescence activated cell sorting (FACS).

3) Fire Fly Luciferase. Firefly luciferase is a 62 kDa protein which catalyzes oxidation of the heterocycle luciferin. The product possesses one of the highest quantum yields for bioluminescent reactions: one photon is emitted for every oxidized luciferin molecule. The structure of luciferase has recently been solved, allowing for structure-based development of a PCA. As with our GST assay, cells are grown on a nitrocellulose matrix. The addition of the luciferin at the surface of the nitrocellulose permits it to diffuse across the cytoplasmic membranes and trigger the photoluminescent reaction. The detection is done immediately on a photographic film. Luciferase is an ideal candidate for a PCA: the detection assays are rapid, inexpensive, very sensitive, and utilizes non-radioactive substrate that is available commercially. The substrate of luciferase, luciferin, can diffuse across the cytoplasmic membrane (under acidic pH), allowing the detection of luciferase in intact cells. This enzyme is currently utilized as a reporter gene in a variety of expression systems. The expression of this protein has been well characterized in

bacterial, mammalian, and in plant cells, suggesting that it would provide a versatile PCA.

4) Xanthine-guanine phosphoribosyl transferase (XGPRT) The *E. coli* enzyme XGPRT converts xanthine to xanthine monophosphate (XMP), a precursor of GMP. Because the mammalian enzyme hypoxanthine-guanine phosphoribosyl transferase HGPRT can only use hypoxanthine and guanine as substrates, the bacterial XGPRT can be used as a dominant selection assay for a PCA for cells grown in the presense of xanthine. Vectors expressing XGPRT confer the ability of mammalian cells to grow in selective medium containing adenine, xanthine, and mycophenolic acid. The function of mycophenolic acid is to inhibit *de novo* synthesis of GMP by blocking the conversion of IMP into XMP (Chapman A. B., (1983) *Molec. & Cellul. Biol.* 3, 1421-1429). The only GMP produced then come from the conversion of xanthine into XMP, catalyzed by the bacterial XGPRT. As with aminoglycoside phosphotransferase fragments of XGPRT can be generated based on the known structure (See table 1.) using the design-evolution strategy described above with fragments fused to the GCN4 leucine zippers as a test oligomerization domains. The complementary fusions are cotransfected and the proteins transiently expressed in COS-7 cells, or stability expressed in CHO cells, grown in the selective medium. In the case of CHO cells, colonies are collected and sequentially re-cultured at increasing concentrations of the selective compounds in order to enrich for populations of cells that efficiently express the fusions at high concentrations.

5) **Adenosine deaminase** Adenosine deaminase (ADA) is present in minute quantities in virtually all mammalian cell. Although it is not an essential enzyme for cell growth, ADA can be used in a dominant selection assay. It is possible to establish growth conditions in which the cells require ADA to survive. ADA catalyzes the irreversible conversion of cytotoxic adenine nucleosides to their respective nontoxic inosine analogues. By adding cytotoxic concentrations of adenosine or cytotoxic adenosine analogues such as 9-b-D-xylofuranosyladenine to the cells, ADA is required for cell growth to detoxify the cytotoxic agent. Cells that incorporate the ADA gene can then be selected for amplification in the presence of low concentrations of 2 \ddot{O} -deoxycoformycin, a tight-binding transition state analogue inhibitor of ADA. ADA can then be used for a PCA based on cell survival (Kaufman, R. J. et al. (1986) Proc. of the Nat. Acad. Sci. (USA) 83, 3136-3140). As with the other systems described above, fragments of ADA can be generated based on the known structure. (See table 1.) using the design-evolution strategy described above with fragments fused to the GCN4 leucine zippers as a test oligomerization domains. The complementary fusions are cotransfected and the proteins transiently expressed in COS-7 cells, or stability expressed in CHO cells, grown in the selective medium containing 2 \ddot{O} -deoxycoformycin. In the case of CHO cells, colonies are collected and sequentially re-cultured at increasing concentrations of 2 \ddot{O} -deoxycoformycin in order to enrich for populations of cells that efficiently express the fusions at high concentrations

25

6) **Bleomycin binding protein (zeocin resistance gene)** Zeocin, a member of the bleomycin/phleomycin family of antibiotics, is toxic to

bacteria, fungi, plants, and mammalian cells. The expression of the zeocin resistance gene confers resistance to bleomycin/zeocin. The protein confers resistance by binding to and sequestering the drug and thus preventing its association and hydrolysis of DNA. Berdy, J. (1980) In *Amino Acid and Peptide Antibiotics*, J. Berdy, ed. (Boca Raton, FL: CRC Press), pp.459-497; Mulsant, P., Tiraby, G., Kallerhoff, J., and Perret, J. (1989 *Somat. Cell. Mol. Genet.* 14, 243-252). Bleomycin binding protein (BBP) could then be used for a PCA based on cell survival. As with the other systems described above, fragments of ADA can be generated based on the known structure (See table 1.) using the design-evolution strategy described above with fragments fused to the GCN4 leucine zippers as a test oligomerization domains. The BBP is a small (8 kD) dimer that binds to drugs *via* a subunit interface binding site. For this reason, the design would be somewhat different in that first, a single chain form of the dimer would be generated by making a fusion of two BBP genes with a short sequence coding for a simple polypeptide linker introduced between the two subunits. Fragments in this case will be based on a short sequence of one of the subunit modules, while the other fragment will be composed of the remaining sequence of the subunit plus the other subunit. Complementation and selection experiments will be performed as described for the examples above using bleomycin or zeocin as selective drugs.

7) Hygromycin-B-phosphotransferase The antibiotic hygromycin-B is an aminocyclitol that inhibits protein synthesis by disrupting translocation and promoting misreading. The *E. coli* enzyme hygromycin-B-phosphotransferase detoxifies the cells by phosphorylating Hygromycin-

B. When expressed in mammalian cells, hygromycin-B-phosphotransferase can confer resistance to hygromycin-B (Gritz, L., and Davies, J. (1983) Gene 25, 179-188.). The enzyme is a dominant selectable marker and could be used for a PCA based on cell survival.

5 While the structure of the enzyme is not known it is suspected that this enzyme is homologous to aminoglycoside kinase (Shaw, et al. (1993) Microbiol. Rev. 57, 138-163). It is therefore possible to use the combined design/evolution strategy to produce a PCA with this enzyme and perform dominant selection in mammalian cells with selection at increasing

10 concentrations of hygromycin B.

8) L-histidinol NAD⁺oxydoreductase The *hisD* gene of *Salmonella typhimurium* codes for the L-histidinol NAD⁺oxydoreductase that converts histidinol to histidine. Mammalian cells grown in media lacking

15 histidine but containing histidinol can be selected for expression of *hisD* (Hartman, S. C., R. C. Mulligan (1988) Proc. of the Nat. Acad. Sci. (USA) 85, 8047-8051). An additional advantage of using *hisD* in dominant selection is that histidinol is itself toxic, inhibiting the activity of endogenous histidyl-tRNA synthetase. Histidinol is also inexpensive and

20 readily permeates cells. The structure of histidinol NAD⁺oxydoreductase is unknown and so development of a PCA based on this enzyme is based entirely on the exonuclease fragment/evolution strategy.

The following Table list alternative embodiments using other PCA reporters. Abbreviations in Table: Type: D, dominant selection marker; R,

25 recessive selection marker. Structure: four letter codes= Protein Data Bank (PDB) entries; K, known but not deposited in PDB; U, unknown. mono/oligo: M, monomer; D, dimer; tetra, tetramer.

TABLE 1. A list of Other Potential PCA Reporter Candidates

A-Assays based on Dominant or Recessive Selection

	Enzyme	Type	Structure	Size	mono/ oligo	Selection drugs/Conditions
5	DHFR	R/D	many	18kD	M	methotrexate/trimetho prim
	Adenosine deaminase	D/R	1ADD		M	Xyl-A or adenosine, alanosine, and 2'- deoxycorformycin
	Thymidine kinase	D/R	1KIN		D	gangcyclovene, HAT
10	Mutant hypoxanthine- guanine phosphoribosyl transferase	D	1HGM		D	HAT + thymidine kinase
	Thymidylate synthetase	R	1NJE	35kd	M	2 fluorodeoxyuridine
	Xanthine-guanine phosphoribosyl transferase	D	1NUL			mycophenolic acid with limiting xanthine
	Glutamine synthetase	R	2LGS			
15	Asparagine synthetase	R	U			B-aspartyl hydroxamate or albizin
	Puromycin N- acetyltransferase	D	U	23kD	M	puromycin
	Aminoglycoside phosphotransferase	D	K	35kD	M	neomycin, G418, gentamycin
20	Hygromycin B phosphotransferase	D	U		M	hygromycin B
	L-histidinol:NAD+ oxidoreductase	D	U	46kD	M	histidinol
	Bleomycin binding protein	D	K	8kD	D	bleomycin/zeocin
25	Cytosine methyl-transferase	R/D	U			5-Azacytidine (5-aza- CR) and 5-aza-2'- deoxycytidine
	O6-alkylguanine alkyltransferase	D	1ADN			N-methyl-N- nitrosourea
	Glycinamide ribonucleotide transformylase	R	1GRC	23.2 kD	D	dideazatetrahydrofolat e, minus purine
30	Glycinamide ribonucleotide synthetase	R	U	45.9 kD		minus purine

	Enzyme	Type	Structure	Size	mono/ oligo	Selection drugs/Conditions
	Phosphoribosyl- aminoimidazole synthetase	R	U	36.7 kD		minus purine
	Formylglycinamide ribotide amidotransferase	R	U	141. 4kD	M	L-azaserine, 6-diazo- 5-oxo-L-norleucine, minus purine
5	Phosphoribosyl- aminoimidazole carboxylase	R	U	39.5 kD	D	minus purine
	Phosphoribosyl- aminoimidazole carboxamide formyltransferase	R	U	57.3 kD		minus purine
10	Fatty acid synthase	R		272k D	D	cerulenin
	IMP dehydrogenase	R	1AK5	55.4 kD	Tetra	mycophenolic acid

15 ii-Viral Plaque Assays

	Enzyme	Type	Structure	Size	Mono/ Oligo	Selection drugs/Conditions
	Thioredoxin	D	1TDF	34.5kD	D	
	Reverse transcriptase	D	3HVT			
20	Viral protease	D				

B-Cell Death Assays

	Enzyme	Type	Structure	Size	Mono/ Oligo	Selection drugs/Conditions
25	Cysteine protease: papain	D	1STF	38.9kD	M	inhibited by cystatin
	Cysteine protease: caspase	D	1CP3	17kD + 12kD	Hetero D	inhibited by DEVD-aldehyde (can also be used in a fluorimetric or colorimetric assay, <i>in vitro</i>)
	Metalloprotease: carboxypeptidase	D		47.1kD	M	inhibited by methyl-ethyl succinic acid
30	Serine protease: proteinase K	D	1PTK	30.6kD	M	inhibited by serpins

5

Enzyme	Type	Structure	Size	Mono/ Oligo	Selection drugs/Conditions
Aspartic protease: pepsin	D	1PSN	34.5kD	M	inhibited by pepstatin A (can also be used in an fluorimetric assay, <i>in vitro</i>)
Lysozyme	D	many	23.2kD	M	inhibited by N-acetylglucosamine trisaccharide
RNAse	D	many	13.3kD	M	inhibited by RNAse inhibitor
DNAse	D	1DNK	61.6kD	M	inhibited by actin
Phospholipase A2	D	1P2P	13.8kD	M/D	many inhibitors: bromophenacyl bromide, hexadecyl-trifluoroethyl-glycero-phosphomethanol, bromoenol lactone, etc.
Phospholipase C	D	1AH7	28kD	M	many inhibitors: neomycin, chelerythrine, U73122, etc.

10

C-Colorimetric/Fluorimetric Assay

15

20

Enzyme	Structure	Size	Mono/ Oligo	Selection drugs/Conditions
DT-Diaphorase (NAD(P)H-[quinone acceptor] oxidoreductase)	1QRD	26kD	D	NADPH-diaphorase stain, inhibited by dicumarol, Cibacron blue and phenidone Note: can also be used in a cell death assay (+nitrobenzimidazole, for example).
(NAD(P)H-[quinone acceptor] oxidoreductase)-2	isoform of 1QRD	21kD	D	NRH-diaphorase stain, inhibited by pentahydroxyflavone
Thermophilic diaphorase (<i>Bacillus stearothermophilus</i>)		30kD	M	NADH-diaphorase stain
Glutathione-S-transferase	1GNE	26kD other isoform of 28kD	D	production of a yellow product by the conjugation of glutathione with an aromatic substance, chloro dinitrobenzene (CDNB)
Luciferase	1LCI	62kD	M	Fluorometric
Green-fluorescent protein	1EMA	30kD	M	Intrinsic fluorescence

Enzyme	Structure	Size	Mono/ Oligo	Selection drugs/Conditions
Chloramphenicol acetyltransferase	1CLA	25kD	Tri	Fluorimetric: Bodipy chloramphenicol
Uricase		32kD	Tetra	Fluorometric
SEAP (secreted form of human placental alkaline phosphatase)	1AJA		M	CSPD chemiluminescent substrate
B-Glucuronidase	1BHG	71kD	Tetra	Histochemical, fluorometric or spectrophotometric assays using various substrates such as X-GLUC.

D-Heteromeric Enzyme Strategies

Tyrosinase		30kD + 14kD	Hetero M+M	Colorimetric: synthesis of melanin
------------	--	-------------	------------	------------------------------------

EXAMPLE 4

Examples of Variants of PCA to detect multiple protein/protein-dna/protein RNA/protein-drug complexes

To this point specific examples have only been made of applications of PCA to protein-pair interactions. However, it is possible to apply PCA to multiprotein, protein-RNA, protein-DNA or protein-small molecule interactions. There are two general schemes for achieving such systems. Multi-subunit PCA: Two proteins need not interact for a PCA signal to be observed; if a partner protein or protein complex binds to two proteins simultaneously, it is possible to detect such a three protein complex. A multisubunit PCA is conceived with the example of herpes simplex virus thymidine kinase (TK), a homodimer of 40 kD. In this conception, the TK structure contains two well defined domains consisting

of an alpha/beta (residues 1-223) and an alpha-helical domain (224-374). As a test system, we use the Rop1 dimer, a four helix bundle homodimer. The two fragments of TK are extracted by PCR and subcloned into the transient transfection vector pMT3, the first in tandem to the Adenovirus
5 major late promoter, tripartite leader 3' to the first ATG, and the second downstream of a ECMV internal initiation site. Restriction sites previously introduced between the first and the last ATG are subcloned into BamHI/ KpnI and PstI/EcoRI cloning sites downstream of the two ATGs. These are used to subclone PCR-generated fragments of the Rop1 subunits into
10 two different vectors. Subsequently Ltk- cells are cotransfected by lipofection with the two plasmids and colonies of surviving cells are serially selected in medium containing increasing concentrations of HAT (hypoxanthine/ aminopterin/thymidine). Cells that express complementary fragments of TK fused to the four Rop1 will proliferate
15 under this selective pressure, or otherwise die. Specific examples of use of this concept would be in determining constituents of multiprotein complexes that are formed transiently or constitutively in cells.

The utility of PCA is not limited to detecting protein-protein interactions, but can be adapted to detecting interactions of
20 proteins with DNA, RNA, or small molecules. In this conception, two proteins are fused to PCA complementary fragments, but the two proteins do not interact with each other. The interaction must be triggered by a third entity, which can be any molecule that will simultaneously bind to the two proteins or induce an interaction between the two proteins by causing
25 a conformational change in one or both of the partners. Two examples have been demonstrated in our lab using the mDHFR PCA in *E. coli*. In the first case a natural product, the immunosuppressant drug rapamycin,

is used to induce an interaction between its receptor FKBP12 and a partner protein mTOR (mammalian Target of Rapamycin). We detect this by cotransformation of DHFR fragments fused to FKBP or mTOR into *E. coli* grown in the presence or absence of trimethoprim (as described above) and rapamycin (0- 10 nM). We have demonstrated that support of growth as detected by colony formation is completely dependent on the addition of rapamycin, suggesting that the mDHFR PCA is detecting a rapamycin-induced assembly of a FKBP12-mTOR and subsequent reconstitution of DHFR activity. This is one example of a use of the PCA strategy to test for small molecules that can induce interactions between proteins. General applications could be made to therapeutic development, in the form screening small molecule combinatorial compound libraries for molecules that induce interactions between proteins, that may inhibit the activities of either or both of the proteins, or activate specific cellular processes that are initiated by other events, such as growth factor-mediated receptor dimerization. The discovery of such small molecules could lead to the development of orally available drugs for the treatment of a broad spectrum of human diseases.

Another example of an induced interaction we have studied with the DHFR PCA is the interaction of the oncogene GTPase p21 ras and its direct downstream target, the serine/threonine kinase raf. This interaction only occurs when the GTPase is in the GTP-bound form, whereas turnover of GTP to GDP leads to release of the complex. As with the FKBP-mTOR complex, we have demonstrated this induced interaction in *E. coli*. PCA could be used in a general way to study such induced interactions, and to screen for compounds that release or prevent these interactions in pathological states. The ras-raf interaction

itself could be a target of therapeutic intervention. Oncogenic forms of ras consist of mutants that are incapable of turning over GTP and therefore remain continuously associated with activated ras. This leads to a constitutive uncontrolled growth signal that results, in part, in oncogenesis. The identification of compounds that inhibit this process, by PCA, would be of value in broad treatment of cancers. Other examples of multimolecular applications of PCA could include identification of novel DNA or RNA binding proteins. In its simplest conception one uses a known DNA or RNA binding motifs, for instance a retinoic acid receptor zinc finger, or a simple RNA binding protein such as IF-1, respectively. One half of the PCA consists of the DNA or RNA protein binding domain fused to one of the PCA fragments (control fragment). The complementary fragment is fused to a cDNA library. A third entity, the gene coding for a sequence containing an element known to bind to the control protein, and then a second putative or known regulatory element is coded for after this sequence. A test system consists of tat/tar elements that control elongation in transcription/translation of HIV genes. An example application would be identification of tat binding elements that have been proposed to exist in eukaryotic genomes and may regulate genes in the same or similar way to that of HIV genes. (SenGupta D. J. et al. (1996) Proc. Natl. Acad. USA 6, 8496-8501).

EXAMPLE 5

Examples of PCA applications to drug screening: Screening combinatorial libraries of compounds for those that inhibit or induce protein-protein/protein-rna/protein-DNA complexes

A) Drug screening

Screening combinatorial libraries of compounds for those that inhibit or induce protein-protein/protein-rna/protein-DNA complexes. The PCA strategy can be directly applied to identifying potentially therapeutic molecules contained in combinatorial libraries of organic molecules. It is possible to perform high throughput screening of such libraries to screen for compounds that will inhibit or induce protein-protein interactions or protein-DNA/RNA interactions (as discussed above). In addition it is also possible to screen for compounds that inhibit enzymes whose substrates are other proteins DNA, RNA or carbohydrates, as discussed below. In this application, proteins that interact/protein substrate pairs, or control DNA/RNA binding protein-enzyme pairs are fused to PCA complementary fragments and plasmids harboring these pairs are transformed/transfected into a cell, along with any third DNA or RNA element as the case requires. Transformed/transacted cells are grown liquid culture in multiwell plates where each well is inoculated with a single compound from an array of combinatorially synthesized compounds. A readout of a response depends on the effect of a compound. If the compound inhibits a protein interaction, there is a negative response (no PCA signal is the positive response). If the compound induces a protein interaction, the response is a positive PCA signal. Controls for non-specific effects of compounds include: 1) demonstration that the compound does not effect the PCA enzyme itself (test against cells transfected with the wild-type intact enzyme used as the PCA probe) and in the case of a cell survival assay, that the compound is not toxic to the cells that have not been transformed/transfected. As well as providing a high throughput assay for

biological activity of compounds, PCA also offers the advantage over *in vitro* assays that it is a test for cell membrane permeability of active compounds. Specific demonstrated examples of PCA for drug screening in our laboratory include the application of DHFR PCA in *E. coli* to
5 detecting compounds that inhibit therapeutically relevant targets. These include Bax/Bcl2 fkbp12/tor ras/raf, carboxyl terminal dimerization domain of HIV-1 capsid protein, Ikb kinase IKK-1 and IKK-2 dimerization domains (leucine zippers and helix-loop-helix domains). In each case, the two proteins are subcloned 5' upstream of either F[1,2] or F[3] as
10 described above. Plasmids harboring the complementary fragments are cotransformed into BL21 cells. Colonies from minimal medium plates containing IPTG and trimethoprim are picked, and grown in liquid medium under the same selective conditions and frozen stocks made. For a single screening cycle, a priming overnight culture is grown from frozen
15 stocks in LB medium. A selective minimal medium containing trimethoprim, ampicillin, IPTG is aliquated at 25 ml into each well of a 384 well plate. Each well is then inoculated with 1 ul of an individual sample from a compound array (ArQule Inc.) to give a final concentration of 10 uM. Each well is then inoculated with 2 ml of overnight culture and
20 plates are incubated in a specially adapted shaker bath at 37C. At 2 hour intervals, plates are read on an optical absorption spectroscopic plate reader coupled to a PC and spreadsheet software at 600 nm (scattering) for a period of 8 hours. Rates of growth are calculated from individual time readings for each well and compared to a standard curve. A "hit" is
25 defined as a case where an individual compound reduces the rate of growth to less than the 95 % confidence interval based on the standard deviation for growth rates observed in all of the wells within the test plate.

"Near hits" are defined as those cases where growth rates are within the 95 % confidence interval. For each of the hits or near hits, the following controls are then performed: The same experiment is performed with BL21 cells that are transformed with empty vector (and no trimethoprim),
5 with vector harboring the full length mDHFR gene, or with cotransfected cells where protein expression is not induced by IPTG. If in all of these cases the compound has no effect, it can be concluded that it is specifically disrupting the protein-protein interaction being tested. Such validated hits or near hits are then retested to establish a dose-response
10 curve for the individual compound, with concentrations varying from 1 pM up to 1 mM by orders of magnitude of 10. The PCA strategy for compound screening can also be applied in the multiprotein protein-RNA/DNA cases as described above, and can easily be adapted to the DHFR or any other PCA in *E. coli* or in yeast versions of the same PCAs.
15 Such screening can also be applied to enzymes whose targets are other proteins or nucleic acids for known enzyme/substrate pairs or to novel enzyme substrate pairs identified as described below.

Proteins involved in viral integration processes are examples of targets that could be tested for inhibitors using the PCA strategies. Examples for
20 the HIV virus include:

- i) Inhibition of integrase or the transport of the pre-integration complex: protein Ma or vpr.
- 25 ii) Inhibition of the cell cycle in G2 by vpr (interaction by cyclin B) causing induction of apoptosis.

iii) Inhibition of the interaction of gp160 (precursor of the membrane proteins) with furine.

Accessory proteins of HIV as a therapeutic target:

5 i) Vpr: nuclear localizing sequence (target): interaction site of vpr with phosphatasesA.

ii) vif: interaction with vimentin (cytoskeleton associated protein) .

10 ii) Vpu: Degradation of CD4 in the RE mediated by the cytoplasmic tail of Vpu.

iii) nef: Myristoylation signal of Nef.

15 Other general targets for drug screening could include proteins linked neurodegenerative diseases, such as to alpha-synuclein. This protein has been linked to early onset of Parkinson disease and it is present also implicated in in Alzheimer disease. There is also b-amyloid proteins, linked to Alzheimers disease.

20 An example of protein-carbohydrate interactions that would be a target for drug screening includes the selectins that are generally implicated in inflammation. These cell surface glycoproteins are directly involved in diapedesis.

25 A number of tumor supressor genes whos actions are mediated by protein-protein interactions could be screened for potential anti-cancer compounds. These include PTEN, a tumor supressor directly involved in the formation of harmatomas. It is also involved in inherited

breast and thyroid cancer. Other interesting tumor suppressor genes include p53, Rb and BARC1.

EXAMPLE 6

5 Examples of applications the PCA strategy to detect enzyme/substrate interactions

The examples described above are used for identifying novel molecular interactions involving molecules that merely bind to each other. However detecting the substrates of enzymes is also fully
10 compatible with the PCA strategy as shown below:

- i) Enzymes that form tight complexes or with protein substrates or induce efficient PCA fragment assembly or
- 15 ii) Mutant enzymes that bind tightly to substrate but do not undergo product release because of mutations residues involved in nucleophilic attack and/or product release (substrate trapping).

Enzymes may form tight complexes with their substrates ($K_d \sim 1-10$ mM). In these cases PCA may be efficient enough to detect
20 such interactions. However, even if this is not true, PCA may work to detect weaker interactions. Generally, if the rate of catalysis and product release is slower than the rate of folding- reassembly of the PCA complementary fragments, effectively irreversible folding and reconstitution of the PCA reporter activity will have occurred. Therefore,
25 even if the enzyme and substrate are no longer interacting, the PCA signal is detected. Therefore, the detection of novel enzyme substrates using PCA may be possible, independent of effective substrate K_d or rate

of product release. In cases where product release is much faster than PCA fragment assembly/folding and alternative approach is provided by generating "substrate trapping" mutants of the test enzyme. An example of this approach applied to the protein tyrosine phosphatase PTP1B, where substrate trapping mutants have been generated by mutating the nucleophilic aspartate 181 to alanine rendering the enzyme catalytically dead, but capable of forming tight complexes with a known substrate, the EGF receptor and other unknown proteins (Flint, A. J. et al. (1996) Proc. Natl. Acad. USA 94:1680-1685). An application of using PCA to screen for interacting partners of PTP1B is given as follows. We use the aminoglycoside kinase (AK)-based PCA in transiently transfected COS or 293 cells. The substrate trap mutant catalytic domain of PTP1B is fused to *N*-terminal complementary fragment of AK, while a C-terminal fusion of the other AK fragment is made to a cDNA library. Cells are co-transfected with complementary AK pairs and grown in selective concentrations of G418. After 72 hours, colonies of surviving cells are picked and in situ PCR is performed using primers designed to anneal to 3' and 5' flanking regions of the cDNA coding region. PCR amplified products are then 5' sequenced to identify the gene.

Enzyme inhibitors Screening combinatorial libraries of compounds for those that inhibit enzyme-PROTEIN substrate complexes either with:

- i) Enzymes that form tight complexes with protein substrates or
- ii) Mutant enzymes that bind tightly to substrate but do not undergo product release because of the mutation.

EXAMPLE 7

Applications of the PCA strategy to protein engineering/evolution

The PCA strategy can be used to generate peptides or proteins with novel binding properties that may have therapeutic value, as is done today with phage display technology. It is also possible to develop enzymes with novel substrate or physical properties for industrial enzyme development. Two detailed examples of the application of the PCA strategy to these ends are given below, with additional applications listed below.

1) **Selection of high-affinity, heterodimerizing leucine zipper sequences** (J. Pelletier, K. Arndt, A. Plueckthun and S. Michnick, manuscript in preparation). The mDHFR PCA, described above, was used in a scheme for the selection of efficiently heterodimerizing, designed leucine zippers. It has been proposed that the formation of salt bridges between positively and negatively charged residues at complementing OeO and OgO positions is important in stabilizing leucine zipper formation, though this view has been contested. In order to help define the importance of salt-bridge formation at the e and g positions, two leucine zipper libraries were built. Both are based on the GCN4 leucine zipper sequence, but contain sequence information specific to either Jun or Fos zippers in order to create heterodimerizing pairs. As well, the e-1 to e-4 and g-1 to g-4 positions in each library were randomized to code for positively or negatively charged residues, or neutral polar residues. These libraries were amplified by PCR and subcloned into the Z-F[1,2] or Z-F[3] constructs (described above) from which the GCN4 zipper sequences had been removed. The bacterial

mDHFR PCA selection was performed on selective solid media, as described earlier. Colonies were picked and sequenced; sequence analysis reveals that the distribution of charged or neutral residues at e-g pairs is not random, but is biased toward pairing of opposite charges, or pairing of a charged with a neutral residue, rather than same-charge pairing (see figure 7). We reasoned that better zipper pairing should lead to an increase in efficiency of DHFR-fragment complementation, resulting in faster bacterial doubling times (see Table 1 in the mDHFR PCA description), and undertook a selection/enrichment of the novel zippers relative to GCN4, as follows. The designed zipper libraries, expressed as N-terminal fusions to the DHFR F[1,2] or F[3:1114A], were cotransformed, clones were picked, propagated and mixed in selective liquid culture, and the mix was added in a 1:1 000 000 ratio to clone Z-F[1,2] + Z-F[3:1114A] (original GCN4 leucine zippers). The mixture was propagated in selective liquid culture over multiple passages. Restriction analysis shows that within 4 passages, the population of GCN4-expressing bacteria is diminishing relative to the novel zipper sequences (data not shown), indicating that some of the designed zipper-containing clones are propagated at a higher rate than those containing GCN4. Bacteria from later passages were plated on selective medium, and individual clones sequenced to reveal the identity of the most successful designed zipper pairs (data not shown).

2) Application of PCA to enzyme function and design PCA

Development: Adenosine deaminase (ADA) meets all of the criteria for a PCA listed above. ADA is a small (~40 kD), and easily purified monomeric zinc metallo-enzyme and the structure of murine ADA has

been resolved. Several *in vitro* ADA activity assays have been developed, involving UV spectrophotometry and stopped-flow fluorimetry. *E. coli* ADA catalyzes the irreversible conversion of cytotoxic adenine nucleosides to non-toxic inosines .

5 Eukaryotic or prokaryotic cells propagated in the presence of cytotoxic concentrations of adenosine or adenosine analogs require ADA to detoxify these compounds. This is the basis of a dominant-selection strategy used to select for cells expressing a specific gene in mammalian cells. The ADA gene has also been expressed in
10 SF3834 *E. coli* cells which lack a gene coding for endogenous ADA. When the gene coding for ADA is introduced into ADA- bacterial DNA, those cells that express ADA are able to survive high concentrations of added adenosine; those that do not, die . This forms the basis of an *in vivo* ADA activity assay.

15 We chose ADA, principally because it can be used as a dominant selective marker in mammalian and bacterial cells where the gene has been knocked out. The reason we choose dominant selective genes is because in screening for novel protein-protein interactions, particularly testing for interactions of a known protein against a library of
20 millions of independent clones, selection serves to filter for cells that may show a positive response for reasons having nothing to do with a specific protein-protein interaction. We will use three test systems of interacting proteins including leucine zipper-forming sequences, the proteins raf and p21 and the induced oligomerization system, FK506 binding protein
25 (FKBP) and mTOR that interact through the macrocyclic immuno-suppressant compound rapamycin. For all of these systems, we will construct *E. coli* and mammalian transient transfection plasmids and

subclone the test proteins as fusions to ADA complementary fragments. The primary assay will be survival of SF3834 *E. coli* cells that have been transformed with the complementary ADA fragments fused to the test oligomerization proteins in the presense of toxic concentrations of adenosine. We will then purify fusion proteins from colonies of and perform *in vitro* assays of ADA activity as described below. The utility of the ADA PCA as a method to identify novel proteins that interact with a test bait will be performed in mammalian COS-7 and HEK-293T cells transiently transfected with FKBP fused to one of the ADA fragments and the other fragment fused to a cDNA library from normal human spleen containing 10^6 independent clones. As with the *E. coli* assay, cells that survive in a medium containing toxic concentrations of ADA is collected and isolated plasmids will be testd to identify the gene for the interacting protein by PCR amplification and chain propagation-termination techniques.

Structural motifs required for protein function: Determination of the structural elements required for the enzymatic function of ADA are investigated through alteration of the structures of the enzyme fragments. At first, ADA is cut into two separate domains - one responsible for substrate binding (residues 1-210) and one responsible for catalysis (residues 211-352). These separate pieces will be attached to known assembly domains, such as leucine zippers (see example 1 above). Reassembly will restore activity which will be assessed through detailed *in vitro* kinetic analysis of the binding and catalytic properties of the re-assembled enzyme, using UV spectrophotometry and stopped-flow fluorimetry to observe the enzymatic reactions. This system will provide

- another handle on the manipulation of enzyme activity that will afford a powerful tool for enzymatic mechanism study. For example, the difference in the kinetic behaviour of the reassembled enzyme on mixing with the substrate, compared to enzyme reassembled in presence of substrate (where substrate may already be bound by binding domain) will allow sophisticated level of study of importance of binding energy to catalysis. Subsequent point mutations to the functional or assembly domains of the proteins will then allow a very subtle perturbation and detailed quantification of the relationship of binding energy to catalysis.
- 10 This precise control over the structure and assembly of separate functional domains of the enzyme will permit very sophisticated enzymatic structure function studies, the definition of structural motifs and an understanding of their role in catalysis.
- 15 ***Novel protein catalyst design:*** The detailed knowledge of the enzyme mechanism gained through determination of the structural requirements for catalysis will then be exploited through the combination of these functional "building blocks" with the functional motifs responsible for substrate binding and catalysis in other enzymes, allowing the generation
- 20 of novel protein catalysts. For example, the catalytic motif from ADA is modified to a cytidine-binding motif, creating a novel enzyme with potentially useful catalytic properties. The activity of these novel enzymes can easily be assessed through *in vivo* assays similar to that of the PCA system, or through *in vitro* activity assays. Furthermore, the
- 25 detailed mechanistic investigation of the resulting enzymes possible with this system will permit the rational design of each subsequent generation of catalysts.

EXAMPLE 8

Examples of applications the PCA strategy to detect molecular interactions in whole organisms

5 It is a logical extension of the descriptions of PCA applications above to the utility of these techniques in whole model organisms such as drosophila, nematodes, zebra fish and puffer fish, as examples. The sole differences with other listed examples is that vectors used would need to be different (for example retroviral vectors) and that any substrates needed by the PCA would need to be bioavailable, or
10 detection would need to be performed *in situ*.

EXAMPLE 9

Examples of applications of the PCA strategy to Gene Therapy

15 Another important embodiment of the invention is to provide a means and method for gene therapy of mammalian disease. Of particular interest is the use of PCA therapeutic for treatment of cancer. In one embodiment of said PCA gene therapy, a PCA is developed employing fragments (modular protein units) derived from a protein toxin for example: Pseudomonas exotoxin, Diptheria toxin and the plant toxin
20 gelonin, or other like molecules. For therapy of breast cancer for example, first a mammalian, retroviral, adenoviral, or eukaryotic artificial chromosomal (EAC's) genetic construct is prepared that introduces one fragment of the selected toxin under the control of the promoter for expression of the *erbB2* oncogene. Its is well known that the *erbB2*
25 oncogene is overexpressed in breast cancer and adenocarcinoma cells (D. J. Slamon et. al., Science, 1989, 244, 707). The *HER2/neu (c-erbB-2)* proto-oncogene encodes a sub-class 1 185-kDa transmembrane

protein tyrosine kinase growth factor receptor, p185^{HER2}. Also, the human *erbB2* oncogene is located on chromosome 17, region q21 and comprises 4,480 base pairs and p185^{HER2} serves as a receptor for a 30-kDa glycoprotein growth factor secreted by human breast cancer cell lines
5 (R. Lupu et. Al., Science, 1990, 249, 1152).

The transgene is introduced 'in vivo' or 'ex-vivo' into target cells employing methods known by those skilled in the art e.g. homologous recombination to insert transgene into locus of interest via retroviral, adenoviral or EAC's. A second genetic construct comprising a
10 fusion gene containing a target DNA that encodes an interacting protein that interacts with *erbB2* oncogene discovered by the PCA process described in this invention and the "second" fragment of the toxin molecule. This construct is delivered to the patient by methods known in the art for example as shown in U.S. Patent Nos. 5,399,346 and
15 5,585,237 whose entire contents are incorporated by reference herein. Transgene expression of the *erbB2* oncogene-toxin fragment described will now be under the control of the constitutive oncogene promoter. Proliferating tumor cells will thus produce one piece of the toxin attached as a fusion to the *erbB2* oncogene. In the presence of the second
20 genetic construct expressing the PCA discovered interacting *erbB2* oncogene "interacting protein - toxin fragment" construct then: *erbB2* oncogene-toxin fragmentA: interacting protein-toxin fragment B will be created and induce death of target tumor cells through creation of an active toxin through Protein Fragment Complementation and thus provide
25 an efficacious and efficient therapy of said disease.

This can be extended to other diseases and other toxins employing techniques described and embodied in this invention.

EXAMPLE 10**Examples of applications the PCA strategy to detect molecular interactions *in vitro***

Any of the PCA strategies described above could be
5 addapted to *in vitro* detection. Unlike the *in vivo* PCAs however,
detection would be performed with purified PCA fragment-fusion proteins.
Such uses of PCA have the potential for use in diagnostic kits. For
example the test DHFR assay described above where the interacting
domains are FKBP12 and TOR could be used as a diagnostic test for
10 rapamycin concentrations for use in monitoring dosage in patients
treated with this drug.

As shown above, the instant invention provides:

- 1) Allow for the detection of protein-protein interactions *in vivo* or *in vitro*.
- 2) Allow for the detection of protein-protein interactions in appropriate
15 contexts, such as within a specific organism, cell type, cellular
compartment, or organelle.
- 3) Allow for the detection of induced versus constitutive protein-protein
interactions (such as by a cell growth or inhibitory factor).
- 4) To be able to distinguish specific-versus non-specific protein-protein
20 interactions by controlling the sensitivity of the assay.
- 5) Allow for the detection of the kinetics of protein assembly in cells.
- 6) Allow for screening of cDNA libraries for protein-protein interactions.

Further aspects of the invention can be demonstrated
by identifying novel interactions with the enzyme p70S6k, to determine its'
25 regulation and how separate signaling cascades converge on this
enzyme.

The PCA method is particularly useful for detection of the kinetics of protein assembly in cells. The kinetics of protein assembly can be determined using fluorescent protein systems.

5 In a further embodiment of the invention, PCA can be used for drug screening. The techniques of PCA are used to screen for drugs that block specific biochemical pathways in cells allowing for a carefully targeted and controlled method for identifying products that have useful pharmacological properties.

10 Although the present invention has been described herein above by way of preferred embodiments thereof, it can be modified, without departing from the spirit and nature of the subject invention as defined in the appended claims.

REFERENCES

1. Reed, L.J.: Multienzyme Complexes. *Acc. Chem. Res.* 7, 40-46 (1974).
2. Lander, E.S.: The new genomics - global views of biology. *Science* 274, 536-539 (1996).
3. Evangelista, C., Lockshon, D. & Fields, S.: The yeast two-hybrid system - prospects for protein linkage maps. *Trends in Cell Biology* 6, 196-199 (1996).
4. Guarente, L.: Strategies for the identification of interacting proteins. [Review]. *Proc. Natl. Acad. Sci. USA* 90, 1639-41 (1993).
5. Adams, S.R., Harootunian, A.T., Buechler, Y.J., Taylor, S.S. & Tsien, R.Y.: Fluorescence ratio imaging of cyclic AMP in single cells. *Nature* 349, 694-7 (1991).
6. Chien, C.T., Bartel, P.L., Sternglanz, R. & Fields, S.: The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest. *Proc. Natl. Acad. Sci. USA* 88, 9578-82 (1991).
7. Fields, S. & Song, O.: A novel genetic system to detect protein-protein interactions. *Nature* 340, 245-6 (1989).
8. Gyuris, J., Golemis, E., Chertkov, H. & Brent, R.: Cdi1, a human G1 and S phase protein phosphatase that associates with Cdk2. *Cell* 75, 791-803 (1993).
9. Johnsson, N. & Varshavsky, A.: Split ubiquitin as a sensor of protein interactions *in vivo*. *Proc. Natl. Acad. Sci. USA* 91, 10340-4 (1994).
10. Volz, K.W., Matthews, D.A., Alden, R.A., Freer, S.T., Hansch, C., Kaufman, B.T. & Kraut, J.: Crystal structure of avian dihydrofolate

- reductase containing phenyltriazine and NADPH. *J. Biol. Chem.* 257, 2528-36 (1982).
11. Oefner, C., D'Arcy, A. & Winkler, F.K.: Crystal structure of human dihydrofolate reductase complexed with folate. *Eur. J. Biochem.* 5 174, 377-85 (1988).
12. Filman, D.J., Bolin, J.T., Matthews, D.A. & Kraut, J.: Crystal structures of *Escherichia coli* and *Lactobacillus casei* dihydrofolate reductase refined at 1.7 Å resolution. II. Environment of bound NADPH and implications for catalysis. *J. Biol. Chem.* 257, 13663-72 10 (1982).
13. Bystroff, C. & Kraut, J.: Crystal structure of unliganded *Escherichia coli* dihydrofolate reductase. Ligand-induced conformational changes and cooperativity in binding. *Biochemistry* 30, 2227-39 (1991).
- 15 14. Appleman, J.R., Prendergast, N., Delcamp, T.J., Freisheim, J.H. & Blakley, R.L.: Kinetics of the formation and isomerization of methotrexate complexes of recombinant human dihydrofolate reductase. *J. Biol. Chem.* 263, 10304-13 (1988).
- 20 15. Loetscher, P., Pratt, G. & Rechsteiner, M.: The C terminus of mouse ornithine decarboxylase confers rapid degradation on dihydrofolate reductase. Support for the pest hypothesis. *J. Biol. Chem.* 266, 11213-20 (1991).
- 25 16. Kaufman, R.J., Davies, M.V., Pathak, V.K. & Hershey, J.W.: The phosphorylation state of eucaryotic initiation factor 2 alters translational efficiency of specific mRNAs. *Mol. Cell. Biol.* 9, 946-58 (1989).

17. Stammers, D.K., Champness, J.N., Beddell, C.R., Dann, J.G., Eliopoulos, E., Geddes, A.J., Ogg, D. & North, A.C.: The structure of mouse L1210 dihydrofolate reductase-drug complexes and the construction of a model of human enzyme. *FEBS Lett.* 218, 178-84 (1987).
18. Gegg, C.V., Bowers, K.E. & Matthews, C.R. in *Techniques in Protein Chemistry* (eds. Marshak, D.R.) 439-448 (Academic Press, New York, USA, 1996).
19. Perry, K.M., Onuffer, J.J., Gittelman, M.S., Barnat, L. & Matthews, C.R.: Long-range electrostatic interactions can influence the folding, stability, and cooperativity of dihydrofolate reductase. *Biochemistry* 28, 7961-8 (1989).
20. Bullerjahn, A.M. & Freisheim, J.H.: Site-directed deletion mutants of a carboxyl-terminal region of human dihydrofolate reductase. *J. Biol. Chem.* 267, 864-70 (1992).
21. Buchwalder, A., Szadkowski, H. & Kirschner, K.: A fully active variant of dihydrofolate reductase with a circularly permuted sequence. *Biochemistry* 31, 1621-30 (1992).
22. Chen, X., Rambo, R. & Matthews, C.R.: Amino acid replacements can selectively affect the interaction energy of autonomous folding units in the alpha subunit of tryptophan synthase. *Biochemistry* 31, 2219-23 (1992).
23. Prevost, M., Wodak, S.J., Tidor, B. & Karplus, M.: Contribution of the hydrophobic effect to protein stability: analysis based on simulations of the Ile-96----Ala mutation in barnase. *Proc. Natl. Acad. Sci. USA* 88, 10880-4 (1991).

24. Kellis, J., Jr., Nyberg, K., Sali, D. & Fersht, A.R.: Contribution of hydrophobic interactions to protein stability. *Nature* 333, 784-6 (1988).
25. Kellis, J., Jr., Nyberg, K. & Fersht, A.R.: Energetics of complementary side-chain packing in a protein hydrophobic core. *Biochemistry* 28, 4914-22 (1989).
26. Henderson, G.B., Russell, A. & Whiteley, J.M.: A fluorescent derivative of methotrexate as an intracellular marker for dihydrofolate reductase in L1210 cells. *Arch. Biochem. Biophys.* 202, 29-34 (1980).
27. Denzer, A.J., Nabholz, C.E. & Spiess, M.: Transmembrane orientation of signal-anchor proteins is affected by the folding state but not the size of the N-terminal domain. *EMBO J.* 14, 6311-7 (1995).
28. Spiess, M., Schwartz, A.L. & Lodish, H.F.: Sequence of human asialoglycoprotein receptor cDNA. An internal signal sequence for membrane insertion. *J. Biol. Chem.* 260, 1979-82 (1985).
29. Picard, V., Ersdal-Badju, E., Lu, A. & Bock, S.C.: A rapid and efficient one-tube PCR-based mutagenesis technique using Pfu DNA polymerase. *Nucleic Acids Res.* 22, 2587-91 (1994).
30. Kunkel, T.A., Roberts, J.D. & Zakour, R.A.: Rapid and efficient site-specific mutagenesis without phenotypic selection. *Methods Enzymol.* 154, 367-82 (1987).
31. Tartof, K.D. & Hobbs, C.A.: New cloning vectors and techniques for easy and rapid restriction mapping. *Gene* 67, 169-82 (1988).
32. Laemmli, U.K.: Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680-5 (1970).

33. Ellenberger, T.E., Brandl, C.J., Struhl, K. & Harrison, S.C.:The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted alpha helices: crystal structure of the protein-DNA complex. *Cell* 71, 1223-37 (1992).
- 5 34. Cody, V., Luft, J.R., Ciszak, E., Kalman, T.I. & Freisheim, J.H.:Crystal structure determination at 2.3 A of recombinant human dihydrofolate reductase ternary complex with NADPH and methotrexate-gamma-tetrazole. *Anti Cancer Drug Des.* 7, 483-91 (1992).
- 10
101. Reed, L.J.:Multienzyme Complexes. *Acc. Chem. Res.* 7, 40-46 (1974).
102. Guarente, L.:Strategies for the identification of interacting proteins. *Proc. Natl. Acad. Sci. USA* 90, 1639-41 (1993).
- 15 103. Adams, S.R., Harootunian, A.T., Buechler, Y.J., Taylor, S.S. & Tsien, R.Y.:Fluorescence ratio imaging of cyclic AMP in single cells. *Nature* 349, 694-7 (1991).
104. Chien, C.T., Bartel, P.L., Sternglanz, R. & Fields, S.:The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest. *Proc. Natl. Acad. Sci. USA* 88, 9578-82 (1991).
- 20 105. Fields, S. & Song, O.:A novel genetic system to detect protein-protein interactions. *Nature* 340, 245-6 (1989).
- 25 106. Gyuris, J., Golemis, E., Chertkov, H. & Brent, R.:Cdi1, a human G1 and S phase protein phosphatase that associates with Cdk2. *Cell* 75, 791-803 (1993).

107. Guarente, L.: Strategies for the identification of interacting proteins. Proc. Natl. Acad. Sci. USA 90, 1639-41 (1993).
108. Johnsson, N. & Varshavsky, A.: Split ubiquitin as a sensor of protein interactions *in vivo*. Proc. Natl. Acad. Sci. USA 91, 10340-4 (1994).
- 5 109. O'Shea, E.K., Klemm, J.D., Kim, P.S. & Alber, T.: X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. Science 254, 539-44 (1991).
110. Ellenberger, T.E., Brandl, C.J., Struhl, K. & Harrison, S.C.: The GCN4 basic region leucine zipper binds DNA as a dimer of
10 uninterrupted alpha helices: crystal structure of the protein-DNA complex. Cell 71, 1223-37 (1992).
111. Volz, K.W., Matthews, D.A., Alden, R.A., Freer, S.T., Hansch, C., Kaufman, B.T. & Kraut, J.: Crystal structure of avian dihydrofolate reductase containing phenyltriazine and NADPH. J. Biol. Chem.
15 257, 2528-36 (1982).
112. Oefner, C., D'Arcy, A. & Winkler, F.K.: Crystal structure of human dihydrofolate reductase complexed with folate. Eur. J. Biochem. 174, 377-85 (1988).
113. Filman, D.J., Bolin, J.T., Matthews, D.A. & Kraut, J.: Crystal
20 structures of Escherichia coli and Lactobacillus casei dihydrofolate reductase refined at 1.7 Å resolution. II. Environment of bound NADPH and implications for catalysis. J. Biol. Chem. 257, 13663-72 (1982).
114. Bystroff, C. & Kraut, J.: Crystal structure of unliganded Escherichia
25 coli dihydrofolate reductase. Ligand-induced conformational changes and cooperativity in binding. Biochemistry 30, 2227-39 (1991).

115. Jones, B.E. & Matthews, C.R.: Early intermediates in the folding of dihydrofolate reductase from *Escherichia coli* detected by hydrogen exchange and NMR. *Protein Sci.* 4, 167-77 (1995).
116. Jennings, P.A., Finn, B.E., Jones, B.E. & Matthews, C.R.: A
5 reexamination of the folding mechanism of dihydrofolate reductase from *Escherichia coli*: verification and refinement of a four-channel model. *Biochemistry* 32, 3783-9 (1993).
117. Fierke, C.A., Johnson, K.A. & Benkovic, S.J.: Construction and evaluation of the kinetic scheme associated with dihydrofolate
10 reductase from *Escherichia coli*. *Biochemistry* 26, 4085-92 (1987).
118. Andrews, J., Fierke, C.A., Birdsall, B., Ostler, G., Feeney, J., Roberts, G.C. & Benkovic, S.J.: A kinetic study of wild-type and mutant dihydrofolate reductases from *Lactobacillus casei*. *Biochemistry* 28, 5743-50 (1989).
119. Thillet, J., Adams, J.A. & Benkovic, S.J.: The kinetic mechanism of
15 wild-type and mutant mouse dihydrofolate reductases. *Biochemistry* 29, 5195-202 (1990).
120. Hillcoat, B.L., Nixon, P.F. & Blakley, R.L.: Effect of substrate decomposition on the spectrophotometric assay of dihydrofolate
20 reductase. *Anal. Biochem.* 21, 178-89 (1967).
121. Appleman, J.R., Prendergast, N., Delcamp, T.J., Freisheim, J.H. & Blakley, R.L.: Kinetics of the formation and isomerization of methotrexate complexes of recombinant human dihydrofolate reductase. *J. Biol. Chem.* 263, 10304-13 (1988).
122. Grange, T., Kunst, F., Thillet, J., Ribadeau-Dumas, B., Mousseron, S., Hung, A., Jami, J. & Pictet, R.: Expression of the mouse dihydrofolate reductase cDNA in *B. subtilis*: a system to select

- mutant cDNAs coding for methotrexate resistant enzymes. *Nucleic Acids Res.* 12, 3585-601 (1984).
123. Loetscher, P., Pratt, G. & Rechsteiner, M.: The C terminus of mouse ornithine decarboxylase confers rapid degradation on dihydrofolate reductase. Support for the pest hypothesis. *J. Biol. Chem.* 266, 11213-20 (1991).
124. Hao, H., Tyshenko, M.G. & Walker, V.K.: Dihydrofolate reductase of *Drosophila*. Cloning and expression of a gene with a rare transcript. *J. Biol. Chem.* 269, 15179-85 (1994).
125. Kaufman, R.J.: Identification of the components necessary for adenovirus translational control and their utilization in cDNA expression vectors. *Proc. Natl. Acad. Sci. USA* 82, 689-93 (1985).
126. Kaufman, R.J., Davies, M.V., Pathak, V.K. & Hershey, J.W.: The phosphorylation state of eucaryotic initiation factor 2 alters translational efficiency of specific mRNAs. *Mol. Cell. Biol.* 9, 946-58 (1989).
127. Oefner, C., D'Arcy, A. & Winkler, F.K.: Crystal structure of human dihydrofolate reductase complexed with folate. *Eur. J. Biochem.* 174, 377-85 (1988).
128. Stammers, D.K., Champness, J.N., Beddell, C.R., Dann, J.G., Eliopoulos, E., Geddes, A.J., Ogg, D. & North, A.C.: The structure of mouse L1210 dihydrofolate reductase-drug complexes and the construction of a model of human enzyme. *FEBS Lett.* 218, 178-84 (1987).
129. Gegg, C.V., Bowers, K.E. & Matthews, C.R. in *Techniques in Protein Chemistry* (eds. Marshak, D.R.) 439-448 (Academic Press, New York, USA, 1996).

130. Bystroff, C. & Kraut, J.: Crystal structure of unliganded *Escherichia coli* dihydrofolate reductase. Ligand-induced conformational changes and cooperativity in binding. *Biochemistry* 30, 2227-39 (1991).
- 5 131. Harmann, B. & Kilimann, M.W.: cDNA encoding a 59 kDa homolog of ribosomal protein S6 kinase from rabbit liver. *FEBS Lett.* 273, 248-52 (1990).
- 10 132. Kozma, S.C., Ferrari, S., Bassand, P., Siegmann, M., Totty, N. & Thomas, G.: Cloning of the mitogen-activated S6 kinase from rat liver reveals an enzyme of the second messenger subfamily. *Proc. Natl. Acad. Sci. USA* 87, 7365-9 (1990).
- 15 133. Grove, J.R., Banerjee, P., Balasubramanyam, A., Coffey, P.J., Price, D.J., Avruch, J. & Woodgett, J.R.: Cloning and expression of two human p70 S6 kinase polypeptides differing only at their amino termini. *Mol. Cell. Biol.* 11, 5541-50 (1991).
134. Banerjee, P., Ahmad, M.F., Grove, J.R., Kozlosky, C., Price, D.J. & Avruch, J.: Molecular structure of a major insulin/mitogen-activated 70-kDa S6 protein kinase. *Proc. Natl. Acad. Sci. USA* 87, 8550-4 (1990).
- 20 135. Reinhard, C., Thomas, G. & Kozma, S.C.: A single gene encodes two isoforms of the p70 S6 kinase: activation upon mitogenic stimulation. *Proc. Natl. Acad. Sci. USA* 89, 4052-6 (1992).
- 25 136. Reinhard, C., Fernandez, A., Lamb, N.J. & Thomas, G.: Nuclear localization of p85s6k: functional requirement for entry into S phase. *EMBO J.* 13, 1557-65 (1994).

137. Chung, J., Kuo, C.J., Crabtree, G.R. & Blenis, J.:Rapamycin-FKBP specifically blocks growth-dependent activation of and signaling by the 70 kd S6 protein kinases. *Cell* 69, 1227-36 (1992).
138. Price, D.J., Grove, J.R., Calvo, V., Avruch, J. & Bierer, B.E.:Rapamycin-induced inhibition of the 70-kilodalton S6 protein kinase. *Science* 257, 973-7 (1992).
139. Kuo, C.J., Chung, J., Fiorentino, D.F., Flanagan, W.M., Blenis, J. & Crabtree, G.R.:Rapamycin selectively inhibits interleukin-2 activation of p70 S6 kinase. *Nature* 358, 70-3 (1992).
140. Terada, N., Franklin, R.A., Lucas, J.J., Blenis, J. & Gelfand, E.W.:Failure of rapamycin to block proliferation once resting cells have entered the cell cycle despite inactivation of p70 S6 kinase. *J. Biol. Chem.* 268, 12062-8 (1993).
141. Calvo, V., Crews, C.M., Vik, T.A. & Bierer, B.E.:Interleukin 2 stimulation of p70 S6 kinase activity is inhibited by the immunosuppressant rapamycin. *Proc. Natl. Acad. Sci. USA* 89, 7571-5 (1992).
142. Weng, Q.P., Andrabi, K., Kozlowski, M.T., Grove, J.R. & Avruch, J.:Multiple independent inputs are required for activation of the p70 S6 kinase. *Mol. Cell. Biol.* 15, 2333-40 (1995).
143. Ming, X.F., Burgering, B.M., Wennstrom, S., Claesson-Welsh, L., Heldin, C.H., Bos, J.L., Kozma, S.C. & Thomas, G.:Activation of p70/p85 S6 kinase by a pathway independent of p21ras [see comments]. *Nature* 371, 426-9 (1994).
144. Cheatham, L., Monfar, M., Chou, M.M. & Blenis, J.:Structural and functional analysis of pp70S6k. *Proc. Natl. Acad. Sci. USA* 92, 11696-700 (1995).

145. Flotow, H. & Thomas, G.: Substrate recognition determinants of the mitogen-activated 70K S6 kinase from rat liver. *J. Biol. Chem.* 267, 3074-8 (1992).
146. Wendt, H., Baici, A. & Bosshard, H.R.: Mechanism of assembly of a leucine zipper domain. *J. Am. Chem. Soc.* 116, 6973-6974 (1994).
147. Mahalingam, M. & Templeton, D.J.: Constitutive activation of S6 kinase by deletion of amino-terminal autoinhibitory and rapamycin sensitivity domains. *Mol. Cell. Biol.* 16, 405-13 (1996).
148. Edery, I., Chu, L.L., Sonenberg, N. & Pelletier, J.: An efficient strategy to isolate full-length cDNAs based on an mRNA cap retention procedure (CAPture). *Mol. Cell. Biol.* 15, 3363-71 (1995).
149. Hussain, A., Lewis, D., Yu, M. & Melera, P.W.: Construction of a dominant selectable marker using a novel dihydrofolate reductase. *Gene* 112, 179-88 (1992).
150. Lim, K., Ho, J.X., Keeling, K., Gilliland, G.L., Ji, X., Ruker, F. & Carter, D.C.: Three-dimensional structure of *Schistosoma japonicum* glutathione S-transferase fused with a six-amino acid conserved neutralizing epitope of gp41 from HIV. *Protein Sci.* 3, 2233-44 (1994).
151. Habig, W.H., Keen, J.H. & Jakoby, W.B.: Glutathione S-transferase in the formation of cyanide from organic thiocyanates and as an organic nitrate reductase. *Biochemical & Biophysical Research Communications* 64, 501-6 (1975).
152. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W.W. & Prasher, D.C.: Green fluorescent protein as a marker for gene expression. *Science* 263, 802-5 (1994).

153. Cody, C.W., Prasher, D.C., Westler, W.M., Prendergast, F.G. & Ward, W.W.: Chemical structure of the hexapeptide chromophore of the *Aequorea* green-fluorescent protein. *Biochemistry* 32, 1212-8 (1993).
- 5 154. Morin, J.G. & Hastings, J.W.: Energy transfer in a bioluminescent system. *Journal of Cellular Physiology* 77, 313-8 (1971).
155. Morise, H., Shimomura, O., Johnson, F.H. & Winant, J.: Intermolecular energy transfer in the bioluminescent system of *Aequorea*. *Biochemistry* 13, 2656-62 (1974).
- 10 156. Ward, W.W. & Bokman, S.H.: Reversible denaturation of *Aequorea* green-fluorescent protein: physical separation and characterization of the renatured protein. *Biochemistry* 21, 4535-40 (1982).
157. Heim, R., Prasher, D.C. & Tsien, R.Y.: Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proc. Natl. Acad. Sci. USA* 91, 12501-4 (1994).
- 15 158. Ormo, M., Cubitt, A.B., Kallio, K., Gross, L.A., Tsien, R.Y. & Remington, S.J.: Crystal structure of the *aequorea victoria* green fluorescent protein. *Science* 273, 1392-1395 (1996).
159. Youvan, D.C. & Michelbeyerle, M.E.: Structure and fluorescence mechanism of gfp. *Nature Biotechnology* 14, 1219-1220 (1996).
- 20 160. DeLuca, M., Woods, K.: Photographic Detection of luminescence in *Escherichia coli* Containing the Gene for Firefly luciferase. *Analytical Biochemistry* 161, 501-507 (1986).
- 25 161. Conti, E., Franks, N.P., Brick, P.: Crystal structure of firefly luciferase throws light on a superfamily of adenylate-forming enzymes. *Structure* 4, 287-298 (1996).

162. Magrath, I.T., Gupta, A., Jain, V.K.:A Rapid Screening Method for Bacteria Contaning Firefly Luciferase Plasmids. Biotechniques 15, 4-6 (1993).

- 5 Abbreviations: PCA, Protein-fragment Complementation Assay; mDHFR, murine dihydrofolate reductase; hDHFR, human dihydrofolate reductase; Z-F[1,2], GCN4 leucine zipper-mDHFR fragment[1,2]; USPS, ubiquitin-based split-protein sensor; IPTG, isopropyl-b-D-thiogalactopyranoside; PMSF, phenylmethylsulfonyl fluoride; SDS-PAGE, SDS polyacrylamide
- 10 gel electrophoresis.

WHAT IS CLAIMED IS:

1. Molecular fragment complementation assays for the detection of molecular interactions comprising a reassembly of separate
5 fragments of a molecule, wherein reassembly of said fragments is operated by the interaction of molecular domains fused to each fragment of said molecules, and wherein reassembly of the fragments is independent of other molecular processes.
- 10 2. A method for detecting biomolecular interactions said method comprising:
 - (a) selecting an appropriate reporter molecule;
 - (b) effecting fragmentation of said reporter molecule such that said fragmentation results in reversible loss of reporter function;
 - 15 (c) fusing or attaching fragments of said reporter molecule separately to other molecules; followed by
 - (d) reassociation of said reporter fragments through interactions of the molecules that are fused to said fragments.
- 20 3. The method of claim 2 wherein said reporter molecule is a multimeric protein.
4. The method of claim 2 wherein said reporter molecule is an multimeric enzyme.
- 25 5. The method of claim 2 wherein said reporter molecule is a multimeric receptor.

6. The method of claim 2 wherein said reporter molecule is a multimeric binding protein.

5 7. The method of claim 2 wherein said reporter molecule is a catalytic molecule.

8. The method of claim 2 wherein said reporter molecule is an energy transfer molecule.

10 9. The method of claim 2 wherein said reporter molecule is a fluorescent or luminescent or phosphorescent protein.

10. The method of claim 2 wherein said detected molecule is a nucleic acid or a ribozyme.

15 11. The method of claim 2 wherein said detected molecule is a lipid or an oligosaccharide.

20 12. The method of claim 2 wherein said detected molecule is a ligand.

13. The method of claim 2 wherein said detected molecule is a nucleic acid.

25 14. The method of claim 2 wherein said detected molecule is a peptide.

15. The method of claim 2 wherein said detected molecule is a carbohydrate.

5 16. The method of claim 2 wherein said fragmentation is effected by a method selected from the group consisting of genetic manipulation, synthetic chemistry or de novo synthesis, photochemical or enzymatic cleavage, and proteolytic or hydrolytic chemistry.

10 17. The method of claim 2 wherein said reassociation of the reporter molecule fragments is effected by molecules fused or attached to said fragments.

18. A method of testing biomolecular interactions comprising:

15 a) generating a first fusion product comprising
i) a first fragment of a first molecule and
ii) a second molecule which is different or the same as said first molecule;

20 b) generating a second fusion product comprising
i) a second fragment of said first molecule; and
ii) a third molecule which is different from or the same as said first molecule or second molecule;

c) allowing the first and second fusion products to contact each other; and

25 d) testing for activity regained by association of the recombined fragments of the first molecule, wherein said reassociation is mediated by interaction of the second and third molecules.

19. The method of claim 18 wherein at least one of said second or said third molecules is a protein.

20. The method of claim 18 wherein at least one of said
5 second or said third molecules is an enzyme.

21. The method of claim 18 wherein at least one of said second or said third molecules is a nucleic acid.

10 22. A method comprising an assay where fragments of a first molecule are fused to a second molecule and fragment association is detected by reconstitution of the first molecule's activity.

23. A composition comprising a product selected from
15 the group consisting of:

(a) a first fusion product comprising:

1) a first fragment of a first molecule whose fragments can exhibit a detectable activity when associated and

2) a second molecule that can bind (a)(1);

20 (b) a second fusion product comprising

1) a second fragment of said first molecule and

2) a third molecule that can bind (b)(1); and

c) both (a) and (b).

25 24. A composition comprising complementary fragments of a first molecule, each fused to separate molecules.

25. The composition of Claim 24 wherein the first molecule is selected from the group consisting of: multimeric protein, multimeric receptor, multimeric binding protein, catalytic molecule, energy transfer molecule, a fluorescent or luminescent or phosphorescent protein.

26. The composition of Claim 24 wherein the first molecule is a multimeric enzyme.

27. The composition of claim 24 wherein the second and third molecules can bind to each other.

28. A composition comprising a nucleic acid molecule coding for a fusion product, which molecule comprises sequences coding for a product selected from the group consisting of:

(a) a first fusion product comprising:

1) fragments of a first molecule whose fragments can exhibit a detectable activity when associated and

2) a second molecule fused to the fragment of the first molecule;

(b) a second fusion product comprising

1) a second fragment of said first molecule and

2) a second or third molecule; and

(c) both (a) and (b).

29. A host cell comprising a composition according to either Claim 24 or Claim 28.

30. An assay which comprises using either the composition according to Claim 24 or Claim 28 or using the host cell according to Claim 29.

- 5 31. A method of testing for biomolecular interactions associated with: (a) complementary fragments of a first molecule whose fragments can exhibit a detectable activity when associated or (b) binding of two protein-protein interacting domains from a second or third molecule, said method comprising:
- 10 1) creating a fusion of
 (a) a first fragment of a first molecule whose fragments can exhibit a detectable activity when associated and
 (b) a first protein-protein interacting domain;
- 15 2) creating a fusion of
 (a) a second fragment of said first molecule and
 (b) a second protein-protein interacting domain
that can bind said first protein-protein interacting domain;
- 3) allowing the fusions of (1) and (2) to contact each other; and
- 20 4) testing for said activity.

32. The method according to any one of Claims 1, 2, 18, 22 or 31 wherein the activity of the fragments of said first molecule is controlled by changing parts of the fragments to either increase or
25 decrease their ability to associate.

33. A composition comprising a product selected from the group consisting of:

- (a) a first fusion product comprising:
 - 1) a first fragment of a molecule whose fragments can exhibit a detectable activity when associated and
 - 2) a first protein-protein interacting domain;
- (b) a second fusion product comprising
 - 1) a second fragment of said first molecule and
 - 2) a second protein-protein interacting domain that can bind said first protein-protein interacting domain; and
- (c) both (a) and (b).

34. A composition comprising a nucleic acid molecule coding for a fusion product, which molecule comprises sequences coding for either:

- (a) a first fusion product comprising:
 - 1) a first fragment of a molecule whose fragments can exhibit a detectable activity when associated and
 - 2) a first protein-protein interacting domain; or
- (b) a second fusion product comprising
 - 1) a second fragment of said molecule and
 - 2) a second protein-protein interacting domain that can bind said first protein-protein interacting domain; or
- (c) both (a) and (b).

35. A host cell comprising a composition according to either Claim 33 or Claim 34.

36. An assay which comprises using either the composition according to Claim 33 or Claim 34 or using the host cell according to Claim 35.

5 37. A method of detecting kinetics of protein assembly comprising performing PCA.

38. A method of detecting kinetics of protein assembly comprising the method of any one of Claims 1, 2, 18, 22, or 31.

10 39. A method of screening a cDNA library comprising performing PCA.

40. A method of screening a cDNA library comprising the method of any one of Claims 1, 2, 18, 22, or 31.

41. A method according to any one of Claims 1, 2, 18, 22, 31, 39 or 40 wherein a chromogenic, fluorogenic, enzymatic, or other optically detectable signal is generated.

20 42. A method according to any one of Claims 1, 2, 18, 22, 31, 39 or 40 wherein dihydrofolate reductase is used as part of a signal-generating system.

25 43. A method of determining the minimum length of at least one of two or more interacting domains comprising performing any of the assays of Claims 1, 2, 18, 22, 31, 39 or 40.

44. A method of determining whether a molecular complex comprises two or more interacting domains, said method comprising performing any of the assays of Claims 1, 2, 18, 22, 31, 39 or 40.

5

45. A method according to any one of Claims 1, 2, 18, 22, 31, 39 or 40 wherein at least one other molecule is present which causes an interaction of said second and third molecules.

10

46. A method according to any one of Claims 39 or 40 where there is an interaction between parts of a first or a second molecule due to the presence of at least one other molecule which mediates the interaction.

15

47. A method of testing the ability of a compound to inhibit molecular interactions in a PCA comprising performing a PCA in the presence of said compound and correlating any inhibition with said presence.

20

48. A method for detecting protein-protein interactions in living organisms and or cells, which method comprises:

(a) synthesizing probe protein fragments from an enzyme which enables dominant selection by dissecting the gene coding for the enzyme into at least two fragments;

25

(b) constructing fusion proteins with one or more molecules that are to be tested for interactions;

(c) fusing the proteins obtained in (b) with one or more of the probe fragments;

(d) coexpressing the fusion proteins; and

(e) detecting the reconstitution of enzyme activity.

5

49. The method according to claim 48 wherein the enzyme is murine dihydrofolate reductase.

50. The method according to claim 48 wherein the
10 fusion proteins have peptides consisting of N- and C-terminal fragments of said murine dihydrofolate reductase and are fused to GCN4 leucine zipper sequences.

51. The method according to claim 48 wherein
15 coexpression of the complementary fusion proteins is catalyzed by the binding of the test proteins to each other.

52. A method for screening or high-throughput
screening of combinatorial libraries for compounds that trigger or inhibit
20 protein-protein interactions, characterized by utilizing the method of claim 48 to identify new drug targets.

53. The method according to claim 52, wherein the
targets are biological active proteins.

25

54. The method of claim 53, wherein said biological active proteins are selected from the group consisting of receptors, inhibitors, enzymes or ion channels.

5 55. A method for detecting biomolecular interactions said method comprising:

- (a) selecting an appropriate reporter molecule;
- (b) effecting fragmentation of said reporter molecule;
- (c) fusing or attaching fragments of said reporter
- 10 molecule separately to other molecules; followed by
- (d) reassociation of said reporter fragments through interactions of the molecules that are fused to said fragments.

56. The composition of Claim 55 wherein the first

15 molecule is selected from the group consisting of: multimeric protein, multimeric receptor, multimeric binding protein, catalytic molecule, energy transfer molecule, a fluorescent or luminescent or phosphorescent protein.

20 57. The composition of Claim 55 wherein the first molecule is multimeric enzyme.

58. In the method of affecting gene therapy, the step which comprises affecting the method of claim 2.

1/7

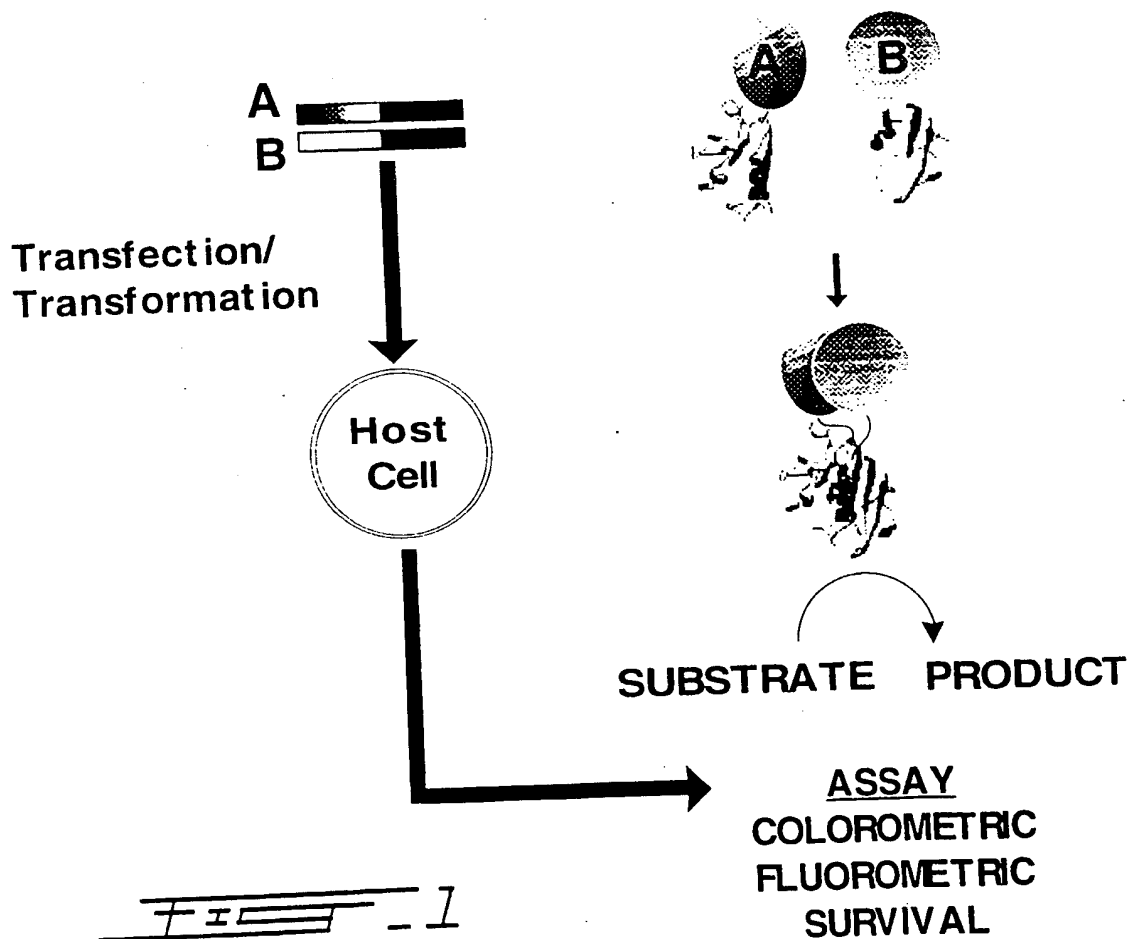
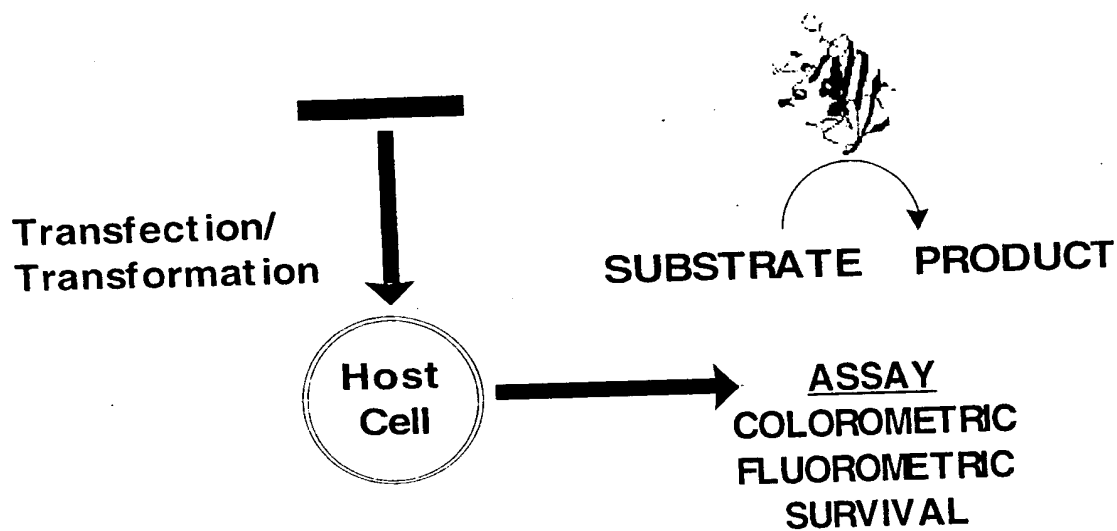


FIG. 1

SUBSTITUTE SHEET (RULE 26)

2/7

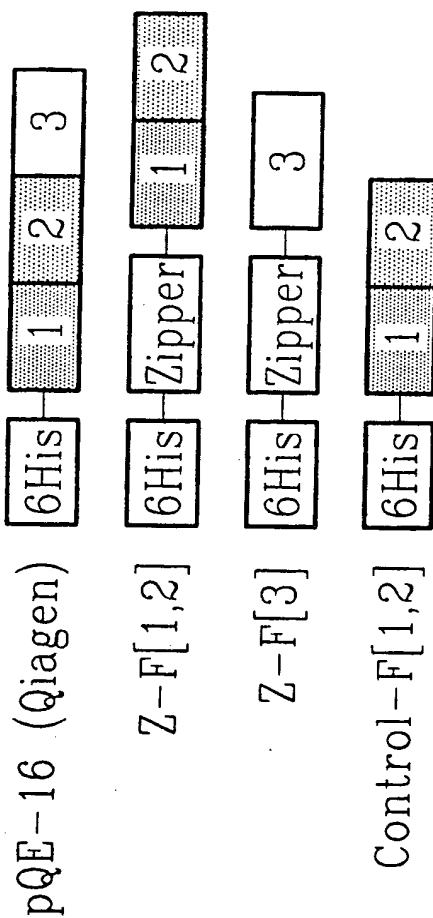
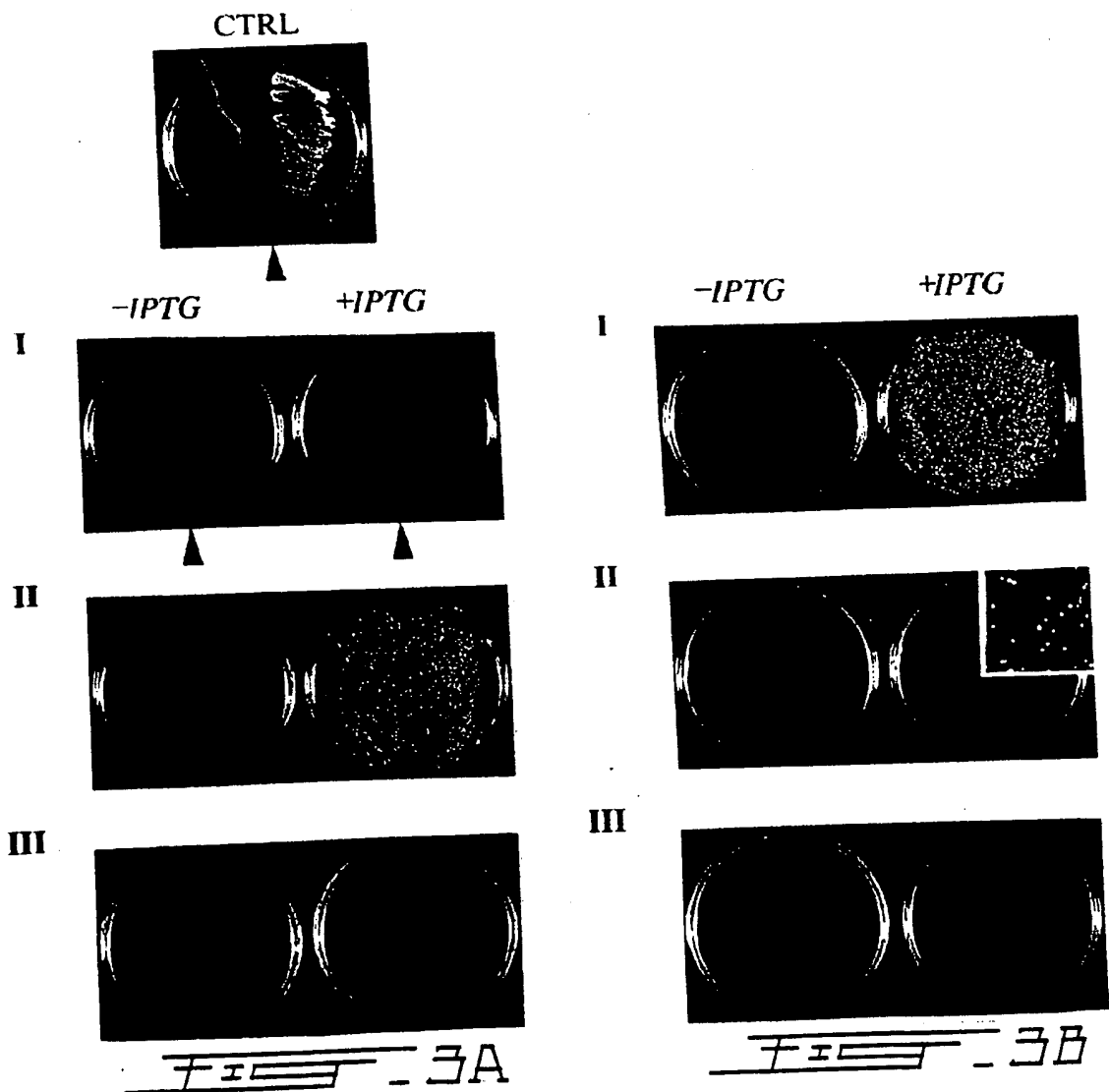


FIG. 2

3/7



SUBSTITUTE SHEET (RULE 26)

4/7

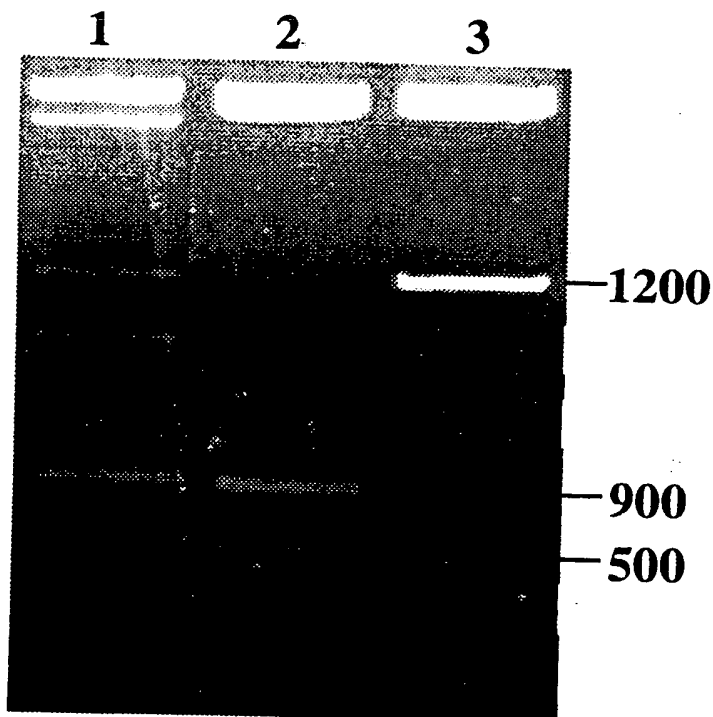


FIG. 4A

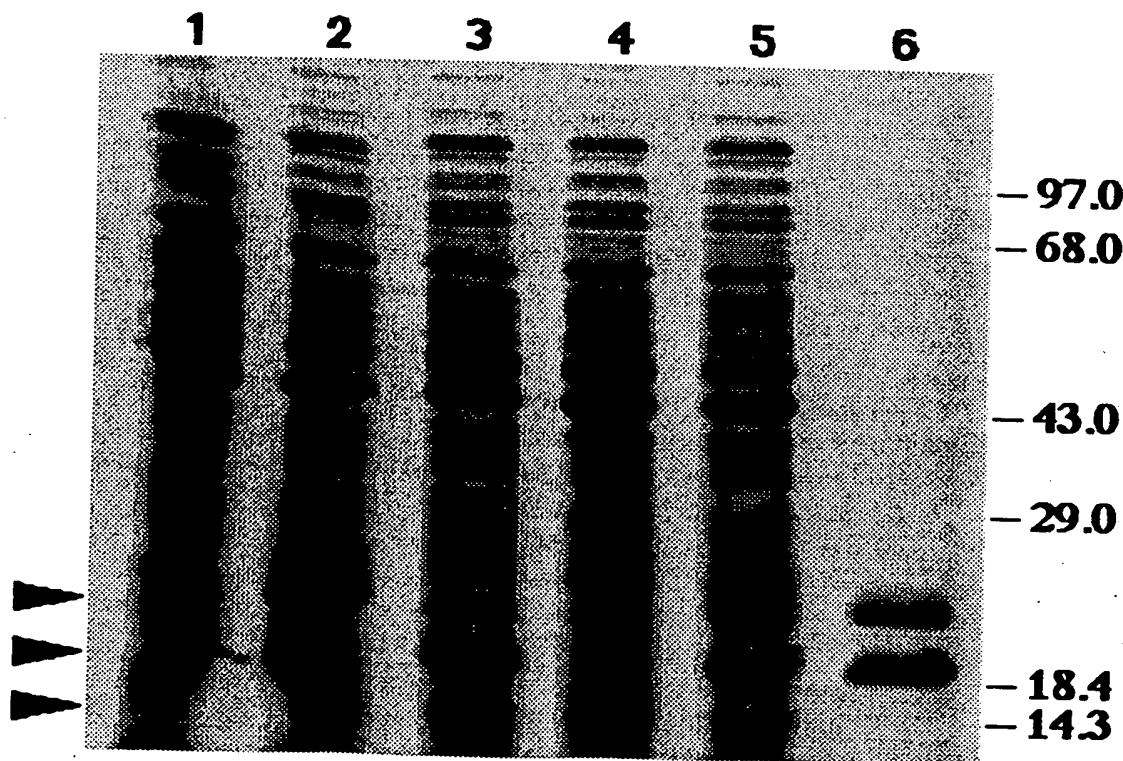
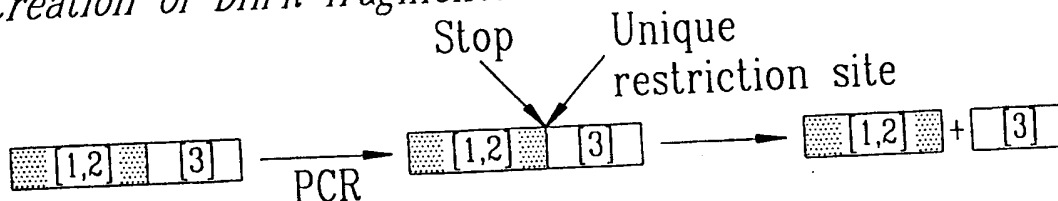


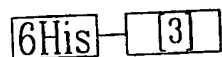
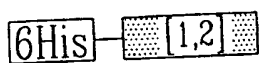
FIG. 4B

SUBSTITUTE SHEET (RULE 26)

5/7

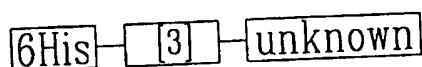
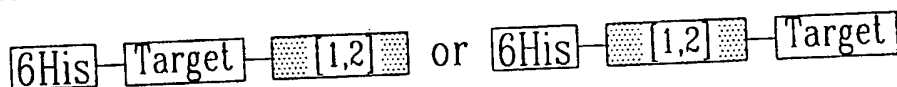
Creation of DHFR fragments:

Insertion of fragments into pQE-32 for bacterial screening, or pMT3 or Zap Express for eukaryotic screening:

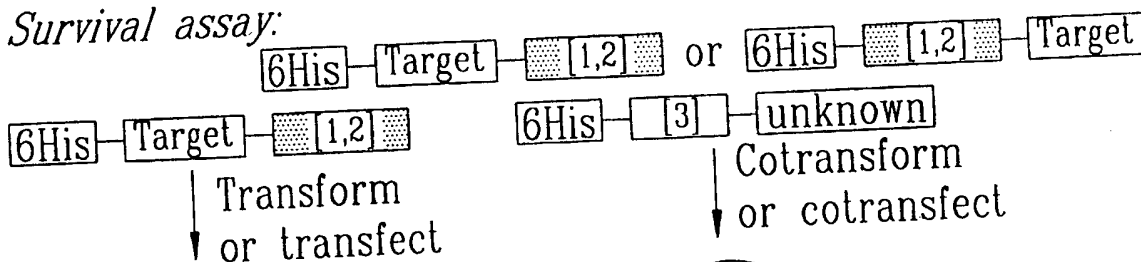


Note: for eukaryotic expression, the expressed constructs have no hexahistidine tag.

Insertion of targets or unknowns (including directional insertion of cDNA library):



Survival assay:



with selective pressure

with selective pressure

Identification of clones (for library screening only):

- Propagation of cells (bacterial or eukaryotic)
- Isolation of plasmid DNA and insert sequencing
- For bacterial screening only: overexpression and one-step purification of fusion products by the hexahistidine tag

FEES 5

6/7

mol. wt. markers
 Zip-[1,2] + Zip-[3]
 Control-[1,2] + Zip-[3]
 Zip-[1,2]
 Zip-[3]
 Control-[1,2]
 wild-type mDHER
 mock

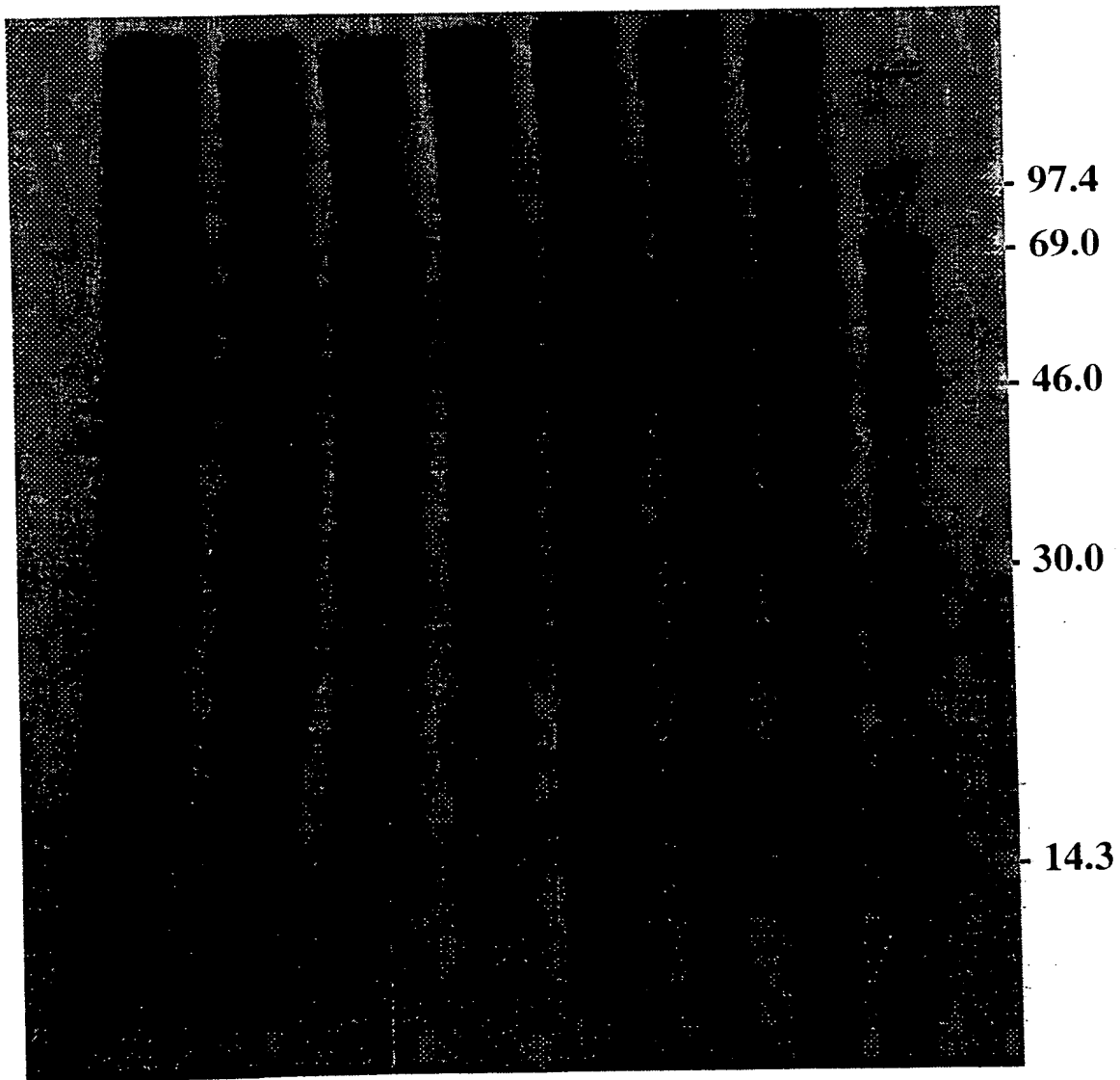
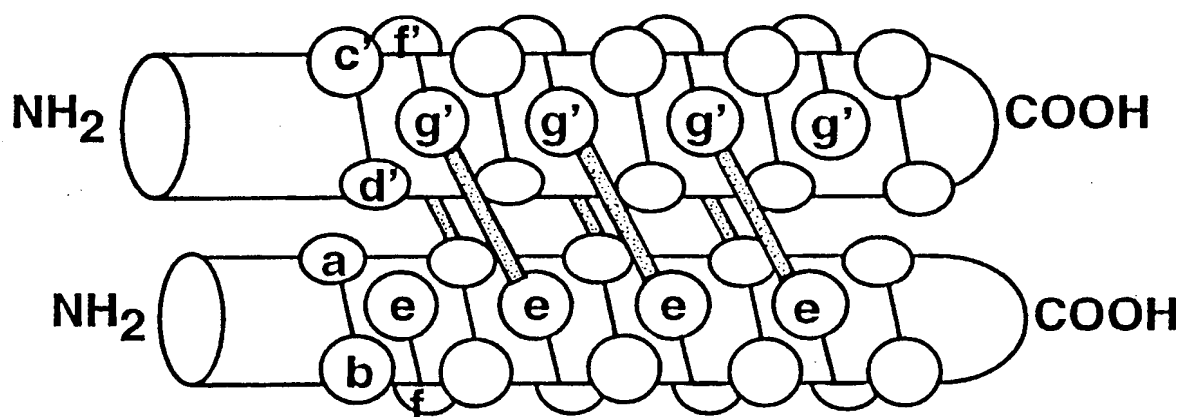
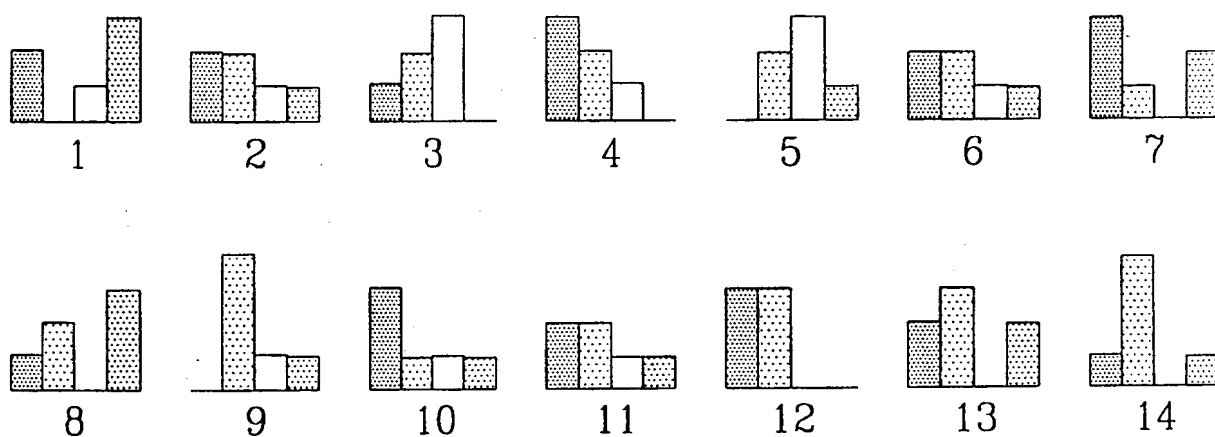


FIG. 6

SUBSTITUTE SHEET (RULE 26)

7/7

FIG. 7A

ATTRACTIVE PAIRING

REPULSIVE PAIRING

■ charge : charge
 ■ charge : neutral polar
 □ neutral polar : neutral polar

■ charge : charge

FIG. 7B

SUBSTITUTE SHEET (RULE 26)

INTERNATIONAL SEARCH REPORT

International Application No

PCT/CA 98/00068

A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 G01N33/68 G01N33/58 G01N33/573

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X, P	J.N. PELLETIER ET AL.: "A protein complementation assay for detection of protein-protein interactions in vivo" PROTEIN ENGINEERING, vol. 10, no. sup, 1 October 1997, OXFORD UK, page 89 XP002064563 see the whole document	1-58
X	US 5 362 625 A (M. KREVON ET AL.) 8 November 1994 cited in the application	1
Y	see the whole document	2-58

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

11 May 1998

Date of mailing of the international search report

26/05/1998

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Van Bohemen, C

INTERNATIONAL SEARCH REPORT

International Application No

PCT/CA 98/00068

C. (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	N. JOHNSON ET AL.: "Split ubiquitin as a sensor of protein interactions in vivo" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES USA, vol. 91, 1 October 1994, WASHINGTON DC USA, pages 10340-10344, XP002064564 cited in the application	1
Y	see the whole document	2-58
X, P	--- F. ROSSI ET AL.: "Monitoring protein-protein interactions in intact eukaryotic cells by beta-galactosidase complementation." PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES USA, vol. 94, 1 October 1997, WASHINGTON DC USA, pages 8405-8410, XP002064565 cited in the application	1
A, P	see the whole document -----	2-58

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/CA 98/00068

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5362625 A	08-11-1994	AT 162891 T	15-02-1998
		AU 657532 B	16-03-1995
		AU 1629192 A	11-03-1993
		CA 2068190 A,C	16-11-1992
		DE 69224231 D	05-03-1998
		EP 0514173 A	19-11-1992
		JP 5276959 A	26-10-1993

